NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

THE AMERICAN WAY OF SWARM: A MACHINE LEARNING STRATEGY FOR TRAINING AUTONOMOUS SYSTEMS

by

Clayton W. Schuety and Lucas E. Will

December 2018

Thesis Advisor:	Robert E. Burks
Co-Advisor:	Michael E. Freeman
Second Reader:	T. Camber Warren

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** December 2018	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
THE AMERICAN WAY OF SWARM: A MACHINE LEARNING STRATEGY FOR TRAINING AUTONOMOUS SYSTEMS

5. FUNDING NUMBERS

6. AUTHOR(S) Clayton W. Schuety and Lucas E. Will

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

Deploying multiple autonomous systems that coordinate as a cohesive swarm on the battlefield is no longer science fiction. As new technologies disrupt the character of war, the American military is investing in algorithms to allow its drone forces to conduct swarm tactics across all domains. However, the current frameworks in development for conducting drone swarm tactics are reliant on centralized control. These frameworks limit the speed and flexibility of the swarm by placing an overreliance on perfect communication and by overtasking the centralized human controller. To overcome these limitations, the American Way of War should adapt; the military must explore novel strategic frameworks that can rapidly train drone algorithms to be effective at decentralized execution, thereby rebalancing the workload of the resulting human-autonomy teams. This thesis proposes that training decentralized swarming algorithms, using the synergy of wargames and machine learning techniques, provides a powerful framework for optimizing drone decision making. The research uses a genetic algorithm to iteratively play a base defense wargame to train local drone interaction rules for a decentralized swarm that generates a desired global behavior. The results show a reduction in average base damage of 78–82% ($p<0.001$) when comparing the mission effectiveness between a pre-trained and a post-trained defensive drone swarm against a baseline adversary.

14. SUBJECT TERMS
autonomous system, swarm, wargame, machine learning, artificial intelligence, training, decision-making, strategy, human-autonomy team, optimization, agent-based model, drone, genetic algorithm, emergent behavior, OODA loop, simulation

15. NUMBER OF PAGES
115

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

THE AMERICAN WAY OF SWARM: A MACHINE LEARNING STRATEGY FOR TRAINING AUTONOMOUS SYSTEMS

Clayton W. Schuety
Lieutenant Colonel, United States Air Force
BS, Iowa State University, 2005
MS, American Military University, 2014

Lucas E. Will
Major, United States Air Force
BS, University of Wisconsin-Madison, 2006
MEng, Pennsylvania State University, 2013

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN DEFENSE ANALYSIS
(IRREGULAR WARFARE)

from the

NAVAL POSTGRADUATE SCHOOL
December 2018

Approved by: Robert E. Burks
Advisor

Michael E. Freeman
Co-Advisor

T. Camber Warren
Second Reader

John J. Arquilla
Chair, Department of Defense Analysis

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Deploying multiple autonomous systems that coordinate as a cohesive swarm on the battlefield is no longer science fiction. As new technologies disrupt the character of war, the American military is investing in algorithms to allow its drone forces to conduct swarm tactics across all domains. However, the current frameworks in development for conducting drone swarm tactics are reliant on centralized control. These frameworks limit the speed and flexibility of the swarm by placing an overreliance on perfect communication and by overtasking the centralized human controller. To overcome these limitations, the American Way of War should adapt; the military must explore novel strategic frameworks that can rapidly train drone algorithms to be effective at decentralized execution, thereby rebalancing the workload of the resulting human-autonomy teams. This thesis proposes that training decentralized swarming algorithms, using the synergy of wargames and machine learning techniques, provides a powerful framework for optimizing drone decision making. The research uses a genetic algorithm to iteratively play a base defense wargame to train local drone interaction rules for a decentralized swarm that generates a desired global behavior. The results show a reduction in average base damage of 78–82% ($p<0.001$) when comparing the mission effectiveness between a pre-trained and a post-trained defensive drone swarm against a baseline adversary.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
AWS	autonomous weapons system
BLUFOR	blue forces
CODA	Coalition for Open-source Defense Analysis
DARPA	Defense Advanced Research Projects Agency
DIUx	Defense Innovation Unit Experimental
DOC	designed operational capability
DoD	Department of Defense
LOCUST	Low-Cost UAV Swarming Technology
MDCOA	most dangerous course of action
MDMP	Military Decision Making Process
REDFOR	red forces
OFFSET	Offensive Swarm-Enabled Tactics
OODA	observe, orient, decide, and act
UAV	unmanned aerial vehicle
UUV	unmanned undersea vehicle
WWII	World War II

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

First and foremost, we would like to thank the Naval Postgraduate School (NPS) and the Defense Analysis department for the educational opportunities here at Monterey, California. The department provided a unique environment to focus on irregular warfare and to develop our critical thinking skills to analyze defense strategy and foreign policy. The diverse academic and operational specialties in the department forced us to challenge our assumptions and to reflect on the complex factors involved in national security. Most importantly, we are grateful for the department's flexible curriculum and willingness to allow for student-led studies that blended operational experiences with the department's diverse faculty perspectives. Ultimately, the flexibility in the curriculum enabled us to pursue a unique research path in autonomous systems and artificial intelligence.

There are countless individuals from across campus who formed an ad-hoc multi-disciplinary team offering their time, insights, and talents to help us refine our approach. Without a doubt, our three advisors were critical in our effort. They pushed us to develop a proof-of-concept model, advancing our ideas beyond merely an untested theory. They also challenged us to pursue a co-evolving wargame, which at first seemed beyond reach. Additionally, we want to thank department professors Kalev Sepp, Leo Blanken, Douglas Borer, and Colonel Michael Richardson for providing their perspectives and mentorship throughout this process. We would also like to thank NPS professors Doug Horner, Duane Davis, Matt Carlyle, and Commander Kathleen Giles, from their diverse areas of expertise, for offering insights that guided our research. Within the student body, we would like to thank Major Nate Padgett, whose previous thesis analyzing swarms using agent-based models inspired our research.

Additionally, we would like to recognize the local support of the Consortium for Robotics and Unmanned Systems Education and Research (CRUSER), which served as a fantastic community to engage with and to build our understanding of current challenges surrounding autonomous systems. In addition, the Modeling, Virtual Environments, and Simulation (MOVES) institute and Coalition for Open-source Defense Analysis (CODA)

labs on campus provided valuable insights and computing resources for the development and execution of our computer-based wargame.

Furthermore, throughout our journey we also met folks outside of the campus who were instrumental in our research. In particular, we would like to thank Doctor Tim Chung at the Defense Advanced Research Projects Agency (DARPA), and Professor Benjamin Jenkins at the Marine Corps University. We would also like to thank Professor Ryan Jenkins from Cal Poly, for making a trip to NPS to offer his perspectives on the critical ethical debates surrounding autonomous systems and their use in national defense.

Finally, but most importantly, we would like to thank our families. Having the opportunity to come to Monterey has been a phenomenal academic endeavor, but it has also been a wonderful chance to spend quality time together as a family, enjoy the local area, and experience many new adventures. We are so grateful for all the love and support you gave us over the course of our academic journey here at NPS. We also want to say thanks for enduring the countless times we dove into conversations at family gatherings that always seemed to come back to “AI” and “swarms”… thanks always, and we love you.

I. INTRODUCTION

> Success no longer goes to the country that develops a new technology first, but rather to the one that better integrates it and swiftly adapts its way of fighting.
>
> —Defense Secretary James Mattis[1]

A. STRATEGIC SITUATION

The 2018 National Defense Strategy assesses that artificial intelligence (AI), machine learning, and autonomous systems will increasingly enable the nation "to gain competitive military advantages."[2] Military researchers also predict that advancements in these technologies will have a compounding effect across all combat domains as more machine autonomy on the battlefield continues to shift the underlying character of war.[3] Most notably, combatants have access to a higher quantity of expendable (yet, capable) autonomous hardware, combined with access to a higher quality of "smart" algorithms.[4] The confluence of these commercially-available technologies is progressing warfare into its next predicted evolution, where any force can now deploy a coordinated collection of autonomous systems (i.e., *swarms*), capable of mounting simultaneous, omnidirectional attacks, into combat.[5]

[1] Colin Clark, "Mattis' Defense Strategy Raises China to Top Threat; Allies Feature Prominently," *Breaking Defense*, January 18, 2018, https://breakingdefense.com/2018/01/mattis-military-strategy-raises-china-to-top-threat-allies-feature-prominently/.

[2] James Mattis, *2018 National Defense Strategy of the United States of America* (Washington, DC: Department of Defense, 2018), 7, https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

[3] Lucia Retter et al., *Moral Component of Cross-Domain Conflict,* RR 1505-MOD (Santa Monica, CA: RAND, 2016), https://www.rand.org/pubs/research_reports/RR1505.html.

[4] Jules Hurst, "Robotic Swarms in Offensive Maneuver," *Joint Force Quarterly*, no. 87 (2017): 105, http://ndupress.ndu.edu/Publications/Article/1326017/robotic-swarms-in-offensive-maneuver/.

[5] John Arquilla and David Ronfeldt, *Swarming and the Future of Conflict, DB 311-OSD* (Santa Monica, CA: RAND, 2000), https://www.rand.org/pubs/documented_briefings/DB311.html; Dave Majumdar, "Who Attacked a Russian Military Base with a 'Swarm' Strike?" *National Interest*, January 12, 2018, https://nationalinterest.org/feature/who-attacked-russian-military-base-swarm-strike-24060.

Already anticipating these "autonomous horizons," the Air Force has plans to integrate more swarming systems into their service by teaming them with the highly trained decision-making skills of their Airmen. However, beyond merely integrating more systems for human operators to control, each service must begin to adapt its way of fighting to delegate more decisions to the trusted algorithms of their autonomous swarms. To improve trust, and thereby retain a competitive advantage, the research of this thesis supports that America's military should adapt a novel *Way of Swarm* that focuses on training not only its people, but training the decision-making algorithms of the swarming autonomous systems, to best optimize the combined human-autonomy teams.

Senior military leaders foreshadow that the emergence of a *general AI* (i.e., a machine with an ability to think, learn, and reason like a human)[6] risks upending the very nature of war.[7] Analysts share this outlook, and agree that the immutable nature of war, often described as a contest of human wills, could eventually be transcended as a contest dominated by AI logic.[8] In 2012, these extreme predictions compelled the Department of Defense (DoD) to focus on general AI investments as a primary effort in its Third Offset strategy that aimed at gaining competitive advantages against adversaries in critical technologies.[9] More recently, the DoD's strategic focus in AI expanded as renewed investments are helping solve a growing "big data" problem.[10] In 2017, the DoD created

[6] Peter Voss, "From Narrow to General AI," *Medium*, October 3, 2017, https://medium.com/intuitionmachine/from-narrow-to-general-ai-e21b568155b9.

[7] James Mattis, "Press Gaggle by Secretary Mattis En Route to Washington, DC," Department of Defense Transcripts, February 17, 2018, https://www.defense.gov/News/Transcripts/Transcript-View/Article/1444921/press-gaggle-by-secretary-mattis-en-route-to-washington-dc/.

[8] "Getting to Grips with Military Robotics: Autonomous Robots and Swarms Will Change the Nature of Warfare," *Economist*, January 25, 2018, https://www.economist.com/special-report/2018/01/25/getting-to-grips-with-military-robotics; Elsa Kania, "Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power," *Center for a New American Security*, November 28, 2017, https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

[9] Paul McLeary, "The Pentagon's Third Offset May Be Dead, But No One Knows What Comes Next," *Foreign Policy*, December 18, 2017, https://foreignpolicy.com/2017/12/18/the-pentagons-third-offset-may-be-dead-but-no-one-knows-what-comes-next/.

[10] Terry Costlow, "How Big Data is Paying off for DoD," *Defense Systems*, October 24, 2014, https://defensesystems.com/articles/2014/10/24/feature-big-data-for-defense.aspx.

a cross-functional team, called Project Maven, that leveraged revolutionary AI programs to help human analysts work through overwhelming amounts of intelligence data.[11] In 2018, the DoD increased its funding for AI further, including $1.75 billion (over 7 years) for the formation of a Joint Artificial Intelligence Center, $93.1 million for Project Maven (a 580% increase from the previous year), $15 million for service-specific AI investments, and over $10 million for the creation of an AI commission.[12] Moreover, in 2018, the DoD announced the removal of the "experimental" status of the Defense Innovation Unit (formerly DIUx) in order to expand and solidify partnerships with the private industries in Silicon Valley that are heavily invested in AI research and development.[13] As a whole, civilian and military investment is pushing AI advancements aggressively in the direction of the *singularity*, defined as the moment when general AI systems will surpass all human intelligence,[14] altering every facet of war (and peace).

To the degree that this singularity becomes reality, a fundamental shift in warfare may prove to be true. Until then, a continued focus on achievements in *narrow AI* (i.e., machine learning techniques)[15] offers a more immediate opportunity to operationalize the emerging technology to tackle complex military problems. One such military problem is the increasing threat of autonomous drone attacks (including swarms), which presents challenges in the air domain that some authors have characterized as the "democratization of airpower."[16] Events in Syria have shown how, for the first time since 1954, American

[11] Cheryl Pellerin, "Project Maven to Deploy Computer Algorithms to War Zone by Year's End," *Department of Defense News*, July 21, 2017, https://www.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

[12] Jay Cassano, "Pentagon's Artificial Intelligence Programs Get Huge Boost in Defense Budget," *Fast Company*, August 15, 2018, https://www.fastcompany.com/90219751/pentagons-artificial-intelligence-programs-get-huge-boost-in-defense-budget.

[13] Aaron Mehta, "Experiment Over: Pentagon's Tech Hub Gets a Vote of Confidence," *Defense News*, August 9, 2018, https://www.defensenews.com/pentagon/2018/08/09/experiment-over-pentagons-tech-hub-gets-a-vote-of-confidence/.

[14] T. C., "What is the Singularity?" *Economist*, May 14, 2018, https://www.economist.com/the-economist-explains/2018/05/14/what-is-the-singularity.

[15] Voss, "From Narrow to General AI," 2017.

[16] T.X. Hammes, "The Democratization of Airpower: The Insurgent and the Drone," *War on the Rocks*, October 18, 2016, https://warontherocks.com/2016/10/the-democratization-of-airpower-the-insurgent-and-the-drone/.

forces have faced threats from the air via adversary drones.[17] Non-state actors, like the Islamic State and Hezbollah, are employing commercial drone technology to contest localized air superiority.[18] Similarly, state actors, like China and Russia, are investing in swarm technology, leveraging the use of higher quantities of cheap systems to offset an American qualitative advantage.[19] With both non-state and state actors advancing drone and swarm technology to gain a competitive advantage, the *American Way of War*[20] risks losing its presumption of air dominance by not adapting to new technologies.

Moreover, as America pivots its National Security Strategy to focus on countering near-peer adversaries, the fight against violent extremist organizations will still require attention and resources.[21] Operating in this complex strategic environment will require innovative solutions that capitalize on the decreasing costs and increasing capabilities of employing trained swarms of autonomous systems. As articulated by Secretary Mattis, technology, alone, will never ensure success in war.[22] Instead, the services need a strategic solution to integrate and operationalize new technologies. The emergence of AI, machine learning, and autonomous systems is calling for a change in the American way of fighting, but overcoming the operational challenges to meet the growing demand for swarming autonomous systems requires an updated framework to better train and equip the emerging human-autonomy teams.

[17] Thomas Gibbons-Neff, “ISIS Drones are Attacking U.S. Troops and Disrupting Airstrikes in Raqqa,” *Washington Post*, June 14, 2017, https://www.washingtonpost.com/news/checkpoint/wp/2017/06/14/isis-drones-are-attacking-u-s-troops-and-disrupting-airstrikes-in-raqqa-officials-say.

[18] Don Rassler, “Remotely Piloted Innovation: Terrorism, Drones, and Supportive Technology,” *Combating Terrorism Center*, October 20, 2016, https://ctc.usma.edu/remotely-piloted-innovation-terrorism-drones-and-supportive-technology/.

[19] Kania, “Battlefield Singularity,” 2017.

[20] Brian McAllister Linn, “The American Way of War Debate: An Overview,” *Historically Speaking* 11, no. 5 (2010): 22–23, https://muse.jhu.edu/article/405440/summary.

[21] Aaron Mehta, “National Defense Strategy Released with Clear Priority: Stay Ahead of Russia and China,” *Defense News*, January 19, 2018, https://www.defensenews.com/breaking-news/2018/01/19/national-defense-strategy-released-with-clear-priority-stay-ahead-of-russia-and-china/.

[22] Clark, “Mattis’ Defense Strategy,” 2018.

B. OPERATIONAL CHALLENGES

The principle challenge to adapting a military force to integrate swarm warfare is solving the span-of-control problem for an increasing quantity of autonomous systems, and thereby rebalancing the workload of the human-autonomy teams.[23] As a recent effort by the Marine Corps Warfighting Lab shows, responsively controlling individual decisions of large numbers of drones is beyond the cognitive capabilities of a single human.[24] The speed, complexity, and scope required to effectively employ a large force of drones will make the current frameworks, which rely heavily on the human operator, obsolete. Therefore, in order to effectively employ drones as a swarm, the human must delegate more freedom of action to the collective decision-making algorithms of the autonomous systems. Delegating more decisions to the autonomous systems enables the human to focus on issues related to mission objectives, risks, and ethical concerns, instead of micromanaging drones' actions in a dynamic battlespace. Although there are efforts focused on making the hardware, software, and interfaces for drones better at a tactical level, there is a gap in the research for how the military can operationalize decentralized mission-specific behaviors for swarms. This gap results in a human-autonomy team that levies significant work on the human to make most of the collective swarming decisions and to centrally control the systems with real-time interfaces.

An additional barrier to operationalizing this emerging technology is the ethical debate that centers on the application of autonomous *weapons* systems (AWS). The use of AWS has received various objections, which include views that their employment generates a responsibility gap, that AWS should not be used to make moral decisions because moral agency is not codifiable, or that even if they were determined to be moral

[23] George Galdorisi, "Keeping Humans in the Loop," *Proceedings* 141, no. 14 (2015), https://www.usni.org/magazines/proceedings/2015-02/keeping-humans-loop; Talya Porat, Tal Oron-Gilad, Michal Rottem-Hovev, and Jacob Silbiger, "Supervising and Controlling Unmanned Systems: A Multi-phase Study with Subject Matter Experts," *Frontiers in Psychology* 7 (2016): 568, https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00568/full.

[24] Gina Harkins, "Marines Test New Drone Swarms a Single Operator Can Control," *Military.com*, July 23, 2018, https://www.military.com/defensetech/2018/07/23/marines-test-new-drone-swarms-single-operator-can-control.html.

agents, they would be making moral decisions based on amoral motivations.[25] In 2017, leaders within the AI community called on the United Nations to ban the development of AWS. Additionally, in 2018, three thousand Google employees demanded an end to the company's partnership with the DoD, stating "We believe that Google should not be in the business of war."[26] However, as technology changes society, society generates defense policy that influences technology; this relationship forms a complex dynamic between war and society.[27] For instance, in 2017, to establish acceptable limits on the development and use of AWS, the civilian-led office for the Under Secretary of Defense for Policy produced a document to guide the services as new commercial technology emerged.[28] Therefore, although the debate is important and ongoing, it does not negate the necessity of utilizing the best technology to counteract threats from adversarial state and non-state actors. As such, this project focuses on objectively improving the strategic frameworks for enhancing swarms, and not focusing on their moral deliberation.

To overcome these operational challenges within the broader strategic situation, this thesis focuses on the following research question: "How can wargames and machine learning be combined to train a decentralized swarm of autonomous systems, thereby enhancing the human-autonomy team?" By implementing a proof-of-concept of the proposed Way of Swarm, the research provides evidence that a machine learning program can repeatedly "self-play" a mission-specific wargame to optimize the decision-making algorithms of an autonomous swarm to (1) achieve the desired overall mission objective and (2) reduce the workload of the human to overcome operational challenges.

[25] Duncan Purves et al., "Autonomous Machines, Moral Judgment, and Acting for the Right Reasons," *Ethical Theory and Moral Practice* 18, no. 4 (2015): 851–872, https://www.researchgate.net/publication/276307723_Autonomous_Machines_Moral_Judgment_and_Acting_for_the_Right_Reasons.

[26] Scott Shane and Daisuke Wakabayashi, "'The Business of War': Google Employees Protest Work for the Pentagon," *New York Times*, April 4, 2018, https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

[27] Rosa Brooks, How Everything Became War and the Military Became Everything: Tales from the Pentagon (New York, NY: Simon and Schuster, 2017).

[28] Department of Defense, *Autonomy in Weapon Systems*, DoD Directive 3000.09 (Washington, DC: DoD, 2017), 36, http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.

C. STRATEGIC FRAMEWORK: TRAINING ALGORITHMS

An examination of factors that have led to the successful employment of military air forces reveals that developing a superior training program (rapid, realistic, and robust) has been decisive in the past.[29] However, as the character of war changes, training better service members is no longer sufficient. In order for autonomous systems to take on more of the decision-making workload, the development of a superior training program is necessary for autonomous algorithms as well. This approach will enhance the human-autonomy team by allowing greater decentralized execution for autonomous systems.

Advancements in AI research have demonstrated that reinforced machine learning techniques are capable of playing against themselves to train algorithms that ultimately outperform the best human minds. Machine learning programs, such as AlphaGo Zero, demonstrated an ability to create novel strategies to complex strategy games, such as Go, which rival the strategies perfected by humans studying the game for generations.[30] In several hours of self-play, the program learned new strategies from an empty slate, unbiased by the constraints of human best practices. AlphaGo Zero now sets a standard for what it means to "train with the best in the world," as machine learning techniques are challenging the understanding for what is possible for a machine to learn and master.

Similarly, this thesis demonstrates that a swarm of autonomous systems could master swarming tactics by combining machine learning principles and reinforced self-play of mission-specific wargames in a rapid, agile, and flexible training framework. Ultimately, the research of this thesis supports that the proposed framework could solve the span-of-control problem in order to better integrate machine learning and autonomous systems into America's military and successfully adapt its way of fighting.

[29] Ralph E. Chatham, "The 20th Century Revolution in Military Training," in *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*, ed. Ericsson KA (UK: Cambridge University Press, 2009), 27–60.

[30] David Silver et al., "Mastering the Game of Go Without Human Knowledge," *Nature* 550, no. 7676 (2017): 354, https://deepmind.com/research/publications/mastering-game-go-without-human-knowledge/.

D. SIGNIFICANCE OF THE RESEARCH

Advances in AI and autonomous systems stand to change the character of war; this necessitates a re-evaluation of the concepts relative to training the operational force. The proposed Way of Swarm demonstrates promise for developing a mastered array of mission-specific tactics for training the DoD's impending swarms of autonomous systems and enhancing the resulting human-autonomy teams. The proposed framework leverages the capabilities of machine learning to self-play through millions of iterations of wargames with a rapid, agile, and flexible process that adjusts to changing real-world assumptions and field-tested observations. Ultimately, training high-quality algorithms for the DoD's drone swarms and teaming them with their service members will help America maintain a competitive advantage... just as a focus on training the best quality service members has been strategically decisive for the military in the past.

The following chapters show the results of thesis research and experimentation, supporting the claim that the proposed framework can effectively train decentralized decision-making algorithms for swarms of autonomous systems. Chapter II begins by exploring the paradigm of swarming warfare, valuable lessons from the study of warfare, and current DoD efforts for developing drone swarm tactics. It concludes by examining advancements in machine learning and its increasing potential for training algorithms. Chapter III outlines the development of a custom-built model, consisting of a swarming wargame and a modified machine learning algorithm, that serves as a proof-of-concept for testing the proposed framework. Chapter IV presents the experimental design and an analysis of the results that evaluates the effectiveness of the custom training model. The analysis shows that in only several days of simulation, the machine learning technique was able to self-play two million iterations of the wargame, improving the success of a base-defense swarm of drones by nearly 80% effectiveness. Finally, Chapter V concludes with major findings from this thesis, relevant applications for the DoD, and the benefits of further research that leverages higher fidelity wargames and more advanced machine learning techniques to operationalize the proposed algorithm training framework.

II. LITERATURE REVIEW

The encountering of swarms of autonomous systems[31] and AI decision-making tools[32] on the battlefield is no longer merely science fiction. Profitable commercial investments are now driving private and academic institutions to research and publish on both of these new technologies.[33] As the National Defense Strategy emphasizes, "The drive to develop new technologies is relentless, expanding to more actors with lower barriers of entry, and moving at an accelerating speed."[34] Consequently, not only are these technologies continuing to disrupt the character of war, but also their associated active research is presenting an opportunity to reassess existing frameworks, and revise them, to maintain a competitive military advantage. By exploring relevant works about these technologies, and by assessing the current DoD-sponsored framework for training swarm tactics, this literature review reveals that the service-wide demand for autonomous systems is outpacing the ability to match inexpensive hardware with quality software for deploying effective autonomous swarms. Therefore, this thesis focuses on an overlooked area in the current literature, investigating how insights from the study of war and recent advancements in machine learning can help address this software gap.

The following chapter is presented in five sections. First, the chapter explores the paradigm of swarming warfare, highlighting the demand and limitations of autonomous systems and their decision-making algorithms to execute decentralized swarm behaviors. Second, the chapter analyzes key lessons learned from the general study of warfare, including why training and mission command principles were critical factors in the past for integrating technology and adapting American forces to a new way of fighting. Third,

[31] Raf Sanchez, "Russia uses Missiles and Cyber Warfare to fight off 'Swarm of Drones' Attacking Military Bases in Syria," *Telegraph*, January 9, 2018, https://www.telegraph.co.uk/news/2018/01/09/russia-fought-swarm-drones-attacking-military-bases-syria/.

[32] Mike Lynch, "AI Cyberattacks Will be Almost Impossible for Humans to Stop," *Wired*, December 28, 2017, https://www.wired.co.uk/article/ai-cyberattack-mike-lynch/.

[33] Daniel Hoadley, *Artificial Intelligence and National Security*, CRS Report No. R45178 (Washington Congressional Research Service, 2018), https://fas.org/sgp/crs/natsec/R45178.pdf.

[34] Mattis, 2018 National Defense Strategy, 2018, 3.

the chapter covers a DoD-sponsored framework for developing swarm tactics and offers an assessment on its strengths and weaknesses. Fourth, the chapter presents research in the field of machine learning, which is revolutionizing algorithm designs for complex strategy games, and demonstrates a potential for training swarming algorithms. Finally, the chapter concludes with how the insight gained improves the proposed Way of Swarm, while also discussing challenges and anticipated critiques to the framework.

A. THE VALUE OF SWARM WARFARE

Employing a formation of relatively simple, replaceable, and independent *agents* that function as members of a formidable force is not a new tactic observed in combat.[35] As with other military tactics, nature first evolved the concepts for employing swarming agents in warfare through millions of years of iterative predator and prey encounters.[36] In natural conflicts, agents of ants, bees, birds, and fish enhanced their species' ability to survive in hostile environments by evolving swarming tactics for defensive and offensive purposes.[37] Mimicking the effectiveness of these biological species, military strategists incorporated similar swarm tactics throughout centuries of warfare. In classical conflicts, agents of horse archers and mounted cavalries executed a variation of swarming tactics, which scholars attribute to the success of leaders such as Genghis Khan and Napoleon.[38] In future conflicts, agents of autonomous systems will continue to improve the strength of these tactics evolved and executed in natural and classical swarms; moreover, they will have the resources to employ swarms at unprecedented scales and speeds.

In the modern context, the term *autonomous system* refers to "a system that can independently compose and select among alternative courses of action to accomplish

[35] Andrew Sanders, "Drone Swarms," (monograph, United States Army Command and General Staff College, 2017), 6–9, http://www.dtic.mil/docs/citations/AD1039921.

[36] Joel Brown and Thomas Vincent, "Organization of Predator-Prey Communities as an Evolutionary Game," *Evolution* 46, no. 5 (1992): 1269–1283, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1992.tb01123.x.

[37] Simon Garnier et al., "The Biological Principles of Swarm Intelligence," *Swarm Intelligence* 1, no. 1 (2007): 3–31, https://link.springer.com/article/10.1007/s11721-007-0004-y.

[38] Sanders, "Drone Swarms," 2017, 6.

goals based on its knowledge and understanding of the world, itself, and the local dynamic context."[39] When referring to a cohesive group of autonomous systems, the term *swarming* characterizes the "collective cooperative dynamics of a large number of decentralized distributed robots through the use of simple local rules."[40] Military scholars contend that swarming tactics composed of autonomous systems will be effective for combat due to a combination of resiliency and versatility.[41] A swarming force is resilient if it has a sufficient quantity of expendable agents to sustain losses, meaning that any particular agent is replaceable and that no single agent is a critical vulnerability to the entire collective. Additionally, a swarming force is versatile if it can best decide when and where to rapidly transition (or "pulse") from executing a dispersed maneuver to establishing a massed presence at a single position. Although a swarm can become more resilient with a greater quantity of autonomous systems, it does not inherently become more versatile, unless each agent can independently make smart local decisions that map to an overall mission objective without a centralized controller.

High-performance and low-cost autonomous systems are becoming the ideal agents for swarming warfare. Commercially available autonomous systems, at the cost of a few hundred dollars each, are capable of automatic flight controls, high-definition imagery, speeds of 45 miles per hour, and altitudes up to 20,000 feet.[42] Modified aerial drone designs can even reach flying speeds of 180 miles per hour.[43] Compared to the cost of a single fighter aircraft ($100M), or military remotely piloted aircraft ($17M), a force

[39] Andrew Ilachinski, "Artificial Intelligence and Autonomy: Opportunities and Challenges," *Center for Naval Analyses*, 2017, https://www.cna.org/cna_files/pdf/DRM-2017-U-014796-Final.pdf.

[40] Ibid.

[41] Kathleen Giles and Kristin Giammarco, "Mission-based Architecture for Swarm Composability (MASC)," *Procedia Computer Science* 114 (2017), https://www.sciencedirect.com/science/article/pii/S1877050917317994.

[42] DJI, "Phantom 4 Specs," 2017, http://www.dji.com/phantom-4/info#specs.

[43] Lisa Segarra, "This Racing Drone Just Set a Guinness World Speed Record," *Fortune*, July 14, 2017, http://fortune.com/2017/07/14/fastest-drone-guinness-world-record/.

could purchase and employ tens of thousands of small aerial drones on the battlefield.[44] With this expendability, losing dozens or hundreds of drones per day becomes an acceptable loss, thereby satisfying the desired swarming trait of resiliency.

Despite improvements in expendable autonomous systems, the methods, and tools necessary for designing quality decision-making software for them to execute versatile swarming tactics continues to lag behind. Research highlights that while "methods exist to facilitate the unique design requirements of robotic swarms, no general method exists that maps individual rules to (desired) group behavior."[45] Thus, there is no current method for leveraging the military's high quantities of low-cost systems with the quality of software needed to be successful. Furthermore, if each drone requires a proportional increase in the workload of human controllers, then there is a diminishing return for fielding the higher quantity swarm. In these cases, success will come down to the quality—not the quantity—of the swarming software.

B. INSIGHT FROM THE PAST

To anticipate what developing quality decision-making software might mean for autonomous swarms, this section examines what developing quality decision-making processes looked like for military forces of the past. Throughout history, the wartime performance of forces has, to a large degree, depended on the quality of their training; it was not the technology that won a battle, but how the overall force operated the hardware that determined the outcome.[46] Historians contend that "the human in the loop is usually the limiting element in the combat effectiveness of the weapon… funding the weapon is not sufficient… we must also fund the warrior."[47] Historical analysis of aerial combat

[44] Christopher Drew, "Lockheed Lowers Price on F-35 Fighters, After Prodding by Trump," *New York Times*, February 3, 2017, https://www.nytimes.com/2017/02/03/business/lockheed-lowers-price-on-f-35-fighters-after-prodding-by-trump.html; Air Force Fact Sheets, "MQ-9 Reaper," September 23, 2015, http://www.af.mil/About-Us/.

[45] Ilachinski, "Artificial Intelligence and Autonomy," 2017, xviii.

[46] Chatham, "The 20th Century Revolution in Military Training," 2009.

[47] Ibid., 59.

provides numerous examples that support this axiom.[48] The Vietnam War provides the most compelling of these examples, driving a revolution in military training across the services. Even though the Air Force led in every aspect of hardware, they suffered heavy losses at the hands of the North Vietnamese Air Force with only a two-to-one air-to-air kill ratio. Although some military historians argue that a number of conflating factors were to blame, others point to the significance of training by citing the success of the Navy's rigorous and realistic Top Gun program, which contributed to a twelve-to-one air-to-air kill ratio during the same timeframe.[49] After the conclusion of Vietnam, both the Army and Air Force built on their experiences and institutionalized the concepts pioneered by the Navy with the development of large force combat exercises like Red Flag and brigade-sized training at the National Training Center.[50]

Institutionalizing realistic exercises, focused on employing hardware at a tactical level, was one way the military leveraged the concept of wargaming to revolutionize its training; but the military also began to leverage computer-based wargames to train operational and strategic level decision making. In this context, *wargames* refer to "analytic games that simulate aspects of warfare at the tactical, operational, or strategic level… used to examine warfighting concepts, train and educate commanders and analysts, [and] explore scenarios."[51] Militaries throughout history have used wargames to gain insight and improve performance dating back to the 5th century B.C. with the Greeks playing *Petteia*, the 6th century A.D. with the Persians and Europeans playing

[48] Anthony H. Cordesman and Abraham R. Wagner, "The Lessons of the 1973 Arab-Israeli Conflict: October War," in *The Lessons of Modern War: Volume 1: The Arab-Israeli Conflict, 1973–1989*, ed. Abraham R. Wagner (Boulder, CO: Westview Press, 1990); Rebecca Grant, "Flying Tiger, Hidden Dragon," *Air Force Magazine*, March 2002, 70–77, http://www.airforcemag.com/MagazineArchive/Documents/2002/March%202002/0302tiger.pdf; Jeffrey S. Johnson, "Initiative in Soviet Air Force Tactics and Decision Making," (master's thesis, Naval Postgraduate School, 1986), https://calhoun.nps.edu/handle/10945/21923; William W. Momyer, "The Counter Air Battle (Air Superiority)," in *Airpower in Three Wars [WWII, Korea, Vietnam]* (Maxwell AFB, AL: Air University Press, 2003).

[49] Brian D. Laslie, *The Air Force Way of War: U.S. Tactics and Training after Vietnam* (Lexington, Kentucky: University Press of Kentucky, 2015).

[50] Jim Robbins, "America's Red Army," *New York Times*, April 17, 1988, https://www.nytimes.com/1988/04/17/magazine/americ-s-red-army.html.

[51] "Wargaming," RAND, accessed September 20, 2018, https://www.rand.org/topics/wargaming.html.

Chess, or the 19th century with Prussians playing *Kriegsspiel*.[52] More recently, the German and Japanese militaries used wargames in WWII before executing their operations in Poland and Pearl Harbor, respectively.[53] Likewise, the United States used wargames in WWII to refine their operational strategy to counter German U-boats.[54] Today, not only are analytic wargames used to provide insight on countering threats like Russia,[55] non-automated wargames are used to "study how people interact in military solutions…[and] study human decision-making."[56] Wargames provide a way to test and develop solutions in a low risk environment[57] that utilize traditional planning methods.[58]

Besides evaluating how forces trained, and integrating computer-based wargames, the military began to focus on ways to maximize the decision-making process of every serviceman. The Air Force started to train aircrews a decision-making model based on *observing* their environment, *orienting* possible solutions, *deciding* based on limited information, and *acting* to achieve a desired effect. Colonel John Boyd, a Korean War fighter pilot, coined this observe, orient, decide, act (OODA loop) decision-making

[52] Roger Smith, "The Long History of Gaming in Military Training," *Simulation & Gaming* no. 41 (2010): 6–19, http://journals.sagepub.com.libproxy.nps.edu/doi/pdf/10.1177/1046878109334330.

[53] Charles Homans, "War Games: A Short History," *Foreign Policy*, August 31, 2011, https://foreignpolicy.com/2011/08/31/war-games-a-short-history/.

[54] William Thomas, "Meta-Calculations and the Mathematics of War," In *Rational Action: The Sciences of Policy in Britain and America, 1940–1960*, ed. Jed Buchwald (Cambridge, MA: MIT Press, 2015), 99–102.

[55] David Shlapak, "The Russian Challenge," PE 250-A (Santa Monica, CA: RAND, 2018), https://www.rand.org/pubs/perspectives/PE250.html

[56] Matthew Schehl and Khaboshi Imbukwa, "Student Wargaming Activities Address Sponsors' Direct Needs," *Naval Postgraduate School*, July 11, 2018, https://my.nps.edu/-/student-wargaming-activities-address-sponsors-direct-needs.

[57] Michael Peck, "Why the Pentagon Loves War Games Again," *National Interest*, May 14, 2016, https://nationalinterest.org/feature/why-the-pentagon-loves-war-games-again-16197; Yuna Huh Wong, "How can Gaming Help Test your Theory?" *RAND Blog*, May 18, 2016, https://www.rand.org/blog/2016/05/how-can-gaming-help-test-your-theory.html.

[58] Department of the Army, *Military Decision-Making Process*, FM 101–5 (Washington, DC: Department of the Army, 1997), http://www.au.af.mil/au/awc/awcgate/army/fm101-5_mdmp.pdf; Department of the Air Force, *Risk Management (RM) Guidelines and Tools*, AF Pamphlet 90–803 (Washington, DC: Department of the Air Force, 2013), http://static.e-publishing.af.mil/production/1/af_se/publication/afpam90-803/afpam90-803.pdf.

process in the 1970s.[59] The Air Force leveraged this concept to "change the way it prepared the aircraft's brain, its pilot, for combat."[60] Boyd taught his OODA loop model at the newly formed Air Force Weapons School under the premise that if a force can make decisions faster than an enemy force, they stand to gain a competitive advantage in combat.[61] Additionally, the Air Force created the designed operational capability (DOC) statement, which leveraged Boyd's concept of accelerating the decision-making cycle.[62] Realizing the cognitive limitations of the human pilot, DOC statements focused each Air Force squadron's OODA loop on only one primary and one secondary mission, as opposed to general utility squadrons that were jacks of all trades, but masters of none. Lastly, the Air Force adopted a building-block training method that initially focused on basic skills and knowledge, gradually built to tactical proficiency, and finally culminated in a validation phase with live complex scenarios that included realistic opposing forces.

Moreover, as the speed, complexity, and scope of military operations increased over the years, the services also managed to succeed in solving their span-of-control limitations by training and leveraging the concept of *decentralized execution*.[63] Historians have credited this concept as a key factor in Napoleon's success, which reshaped warfare in his time, and continues to have relevance today.[64] Specifically, decentralized execution emerged as one of the fundamental tenets of the "Air Force Way of War," as it adapts to new generations of aircraft with exceptional speeds, range, and

[59] Frans Osinga, Science, Strategy, and War: The Strategic Theory of John Boyd (London, UK: Routledge, 2007).

[60] Laslie, Air Force Way of War, 2015.

[61] Robert Coram, *Boyd: The Fighter Pilot Who Changed the Art of War* (New York, NY: Little, Brown and Company, 2002).

[62] Laslie, Air Force Way of War, 2015.

[63] Clint Hinote, *Centralized Control and Decentralized Execution: A Catchphrase in Crisis?* (Maxwell AFB, AL: Air Force Research Institute, 2009), https://permanent.access.gpo.gov/gpo23521/a550460.pdf.

[64] Antoine Bousquet, *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity* (London, UK: Hurst and Company, 2009); Jim Storr, "A Command Philosophy for the Information Age: The Continuing Relevance of Mission Command," *Defense Studies* 3, no. 3 (2003): 119–129, https://www.tandfonline.com/doi/pdf/10.1080/14702430308405081.

flexibility.[65] All of the services now use the term *mission command* to embody the importance of maintaining decentralized execution in war.[66] Mission command benefits forces by preserving tactical flexibility through a commander's intent, allowing subordinate leaders to seize initiative by acting aggressively and thinking independently, to accomplish their assigned mission.

Just as rigorous training, computer-based wargaming, faster decision making, and decentralized execution were critical factors in past contexts, it is reasonable to assume they will be critical in future contexts with swarms of autonomous systems. Researchers looking at future wars contend that the decisive factor in the quality of autonomous swarms will be the OODA-loop-like algorithms inside the hardware that will act either independently or in tandem with the human operators.[67] In short, training humans will no longer be adequate for maintaining a competitive advantage. Instead, training algorithms will become increasingly critical to maximizing the effectiveness of the human-autonomy teams. Quality algorithms will enable more flexibility in the tactical environment and allow for decentralized execution at an unprecedented scale. Thus, as the character of war changes, a strategic framework that can rapidly train algorithms and can build trust and confidence between the human and the autonomous system holds vast potential.[68]

C. FRAMEWORKS FOR TRAINING SWARM TACTICS

The revolution in the American military training programs after the Vietnam War lack a fundamental component in today's changing environment: strategies for not only developing operators to optimize the human decision-making process, but for training the autonomous systems to observe, orient, decide, and act on behalf of human-specified

[65] Laslie, Air Force Way of War, 2015.

[66] Department of the Army, *Operations*, FM 3–0 (Washington, DC: Department of the Army, 2017), https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN6687_FM%203-0%20C1%20Inc%20FINAL%20WEB.pdf.

[67] Peter W. Singer, *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (New York, NY: Penguin Press, 2009).

[68] Ranjeev Mittu, Donald Sofge, Alan Wagner, and William Frere Lawless, *Robust Intelligence and Trust in Autonomous Systems* (Boston, MA: Springer, 2016).

objectives. Autonomy, implemented through refined algorithms, allows machines to process and multitask decision-making loops independently when assigned broad tasks by humans. Technology has advanced to the point where the human decision-making cycle could limit the overall potential of swarms of autonomous systems. Therefore, the DoD should continue to pursue ways to optimize the human-autonomy team that involve delegating more freedom of action to their swarming autonomous systems (Figure 1).[69]

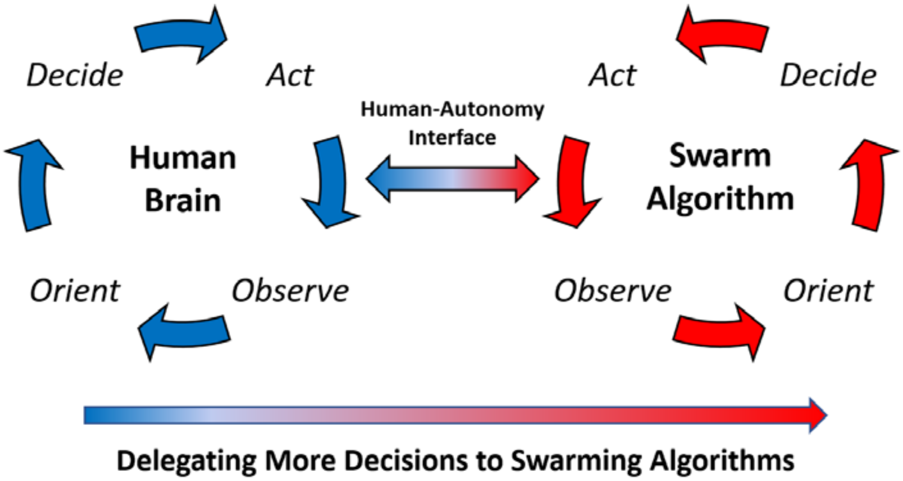


Figure 1. Decision-Making Process for a Human-Autonomy Team[70]

Efforts to enable cooperative dynamics or to engineer active swarming behaviors are still nascent in the military's pursuits for autonomous systems in the air, land, and sea domains.[71] For example, in 2015, the Office of Naval Research, under the Low-Cost UAV Swarming Technology (LOCUST) program, successfully controlled thirty aerial

[69] Department of Defense, *Task Force Report: The Role of Autonomy in DoD Systems* (Washington, DC: Department of Defense, 2012), https://fas.org/irp/agency/dod/dsb/autonomy.pdf.

[70] Adapted from Osinga, *Strategic Theory of John Boyd*, 2007, 2; U.S. Air Force Office of the Chief Scientist, *Autonomous Horizons*, 2015, 7.

[71] Daniel Gonzales and Sarah Harting, *Designing Unmanned Systems with Greater Autonomy*, RR 626-OSD (Santa Monica, CA: RAND, 2014), https://www.rand.org/content/dam/rand/pubs/research_reports/RR600/RR626/RAND_RR626.pdf.

drones autonomously over a predetermined path.[72] In 2016, the program conducted its largest test by airdropping and controlling 103 Perdix micro-drones from an F-18 Super Hornet. The Perdix drones conducted low-altitude reconnaissance missions and operated in “small and large swarms to perform their missions.”[73] Most recently, in 2018, the Navy added its first drone ship to the fleet, with future hopes of interconnecting multiple vessels as an autonomous swarm to scan the world’s oceans.[74] Executing at even larger scales, commercial companies hold the world record for the highest number of drones controlled simultaneously. American and Chinese companies, such as Intel and Ehang, flew a thousand “light-show” quadcopter drones for marketing and entertainment value.[75] Although all of these military and commercial efforts are advancing the field of swarm research and application, they still face major challenges. First, they all require extensive man-hours to develop the swarm algorithms. Additionally, the developed algorithms are “static” in that they are only able to conform to pre-programmed patterns and formations. In other words, the swarming agents lack a robust set of individual rules that drive a dynamic and emergent group behavior. Lastly, they all require significant workload from the human operator during execution to achieve their specific objectives.

Many working in the field of autonomous systems research have recognized the value of training swarms, but they have used different approaches and means to leverage its benefits. Some teams have taken a *bottom-up approach*, looking to develop complex

[72] Kevin McCaney, “Day of the LOCUST: Navy Demonstrates Swarming UAVs,” *Defense Systems*, April 15, 2015, https://defensesystems.com/articles/2015/04/15/onr-locust-swarming-autonomous-uavs.

[73] Department of Defense, “Perdix Fact Sheet,” Accessed on February 11, 2018, https://dod.defense.gov/Portals/1/Documents/pubs/Perdix%20Fact%20Sheet.pdf.

[74] Joseph Trevithick, “Navy’s Sea Hunter Drone Ship Is Getting a New Owner, New Abilities, and a Sister,” *Drive*, February 6, 2018, http://www.thedrive.com/the-war-zone/18264/navys-sea-hunter-drone-ship-is-getting-a-new-owner-new-abilities-and-a-sister.

[75] Brian Barrett, “Inside the Olympics Opening Ceremony World-Record Breaking Drone Show,” *Wired*, February 9, 2018, https://www.wired.com/story/olympics-opening-ceremony-drone-show/; Jeffrey Lin and P.W. Singer, “China is Making 1,000-UAV Drone Swarms Now,” *Popular Science*, January 8, 2018, https://www.popsci.com/china-drone-swarms.

emergent swarm behaviors from basic subcomponents.[76] These teams have used agent-based models to study rule sets created for groups of autonomous agents and how they interact with a simulation environment. Others argue that the bottom-up approach can "often risk failing to meet higher-level system requirements if design begins before a higher-level system architecture is established."[77] Instead, these researchers put forth a *top-down approach* for developing a framework of phases, tactics, plays, and basic algorithms that nest under a specific military mission.

One DoD sponsored effort for training swarms through a top-down approach is DARPA's Offensive Swarm-Enabled Tactics (OFFSET) program.[78] OFFSET proposes to use a real-time game environment, and a virtual reality interface, to allow users to derive novel swarm tactics for autonomous systems through crowd-sourcing methods.[79] By applying a top-down approach to swarm tactic designs,[80] and by using mission-specific games to train, test, and employ swarming capabilities, the OFFSET framework is also helping to advance the field of swarm design. The program plans to pair its framework with baseline swarm characteristics produced by the Navy's LOCUST program.[81]

Despite the advantages of OFFSET, there are three limitations with this current DoD framework. First, relying on crowd-sourcing efforts may be problematic to maintain

[76] Andrei Borshchev and Alexei Filippov, "From System Dynamics and Discrete Event to Practical Agent Based Modeling," In *International Conference of the System Dynamics Society* (July 2004), 25–29, https://www.systemdynamics.org/assets/conferences/2004/SDS_2004/PAPERS/381BORSH.pdf; Ryan McCune, et al., "Investigations of DDDAS for Command and Control of UAV Swarms with Agent-Based Modeling," In *Proceedings of the 2013 Winter Simulation Conference: Making Decisions in a Complex World* (December 2013), 1467–1478, https://ieeexplore.ieee.org/document/6721531; Mauricio Munoz, "Agent-based Simulation and Analysis of a Defensive UAV Swarm Against an Enemy UAV Swarm" (master's thesis, Naval Postgraduate School, 2011), https://calhoun.nps.edu/handle/10945/5700.

[77] Giles and Giammarco, "Mission-based Architecture for Swarm Composability (MASC)," 2017.

[78] "OFFensive Swarm-Enabled Tactics (OFFSET)," DARPA, accessed September 20, 2018, https://www.darpa.mil/work-with-us/offensive-swarm-enabled-tactics.

[79] Alec Meden, "DARPA's Game of Drones," *Atlantic Council*, accessed October 9, 2018, http://artoffuturewarfare.org/2016/12/darpas-game-of-drones/.

[80] Giles and Giammarco, "Mission-based Architecture for Swarm Composability (MASC)," 2017.

[81] "DARPA Adds Two Companies to OFFSET Swarm Reconnaissance Drone Research Project," *Military Aerospace*, May 8, 2018, https://www.militaryaerospace.com/articles/print/volume-29/issue-4/unmanned-vehicles/darpa-adds-two-companies-to-offset-swarm-reconnaissance-drone-research-project.html.

over time; the size of the "crowd" may not be sustainable as interest (and funding) in the project ebbs and flows. Second, as different hardware and environments are ready for testing, a crowd-sourcing reliant method is cumbersome to rapidly repeat in training. The framework would require lengthy real-time replays of all the previously generated data to determine what behaviors are now obsolete or what algorithms may have become better tactical solutions with any changes to the hardware or environmental assumptions. Third, OFFSET over-emphasizes the need for real-time execution of swarms.[82] The framework uses a controlling application where drone swarms move via "point-and-click" through the battlespace. Not only does this approach reduce the speed and initiative of swarms in operations, but the real-time aspect of the application limits the ability to speed up repetitions to train through thousands of potential tactical scenarios in seconds.

Understanding that there is a demand for DoD swarming systems, the importance of training to generate quality decision-making processes, and that there is a gap in the current framework for training quality algorithms, the literature review now pivots to research in the field of narrow AI and machine learning to present a novel solution.

D. REVOLUTIONS IN MACHINE LEARNING

As with the theory behind swarming tactics, the theory behind machine learning is also not conceptually "new." Modeling decision making as an "artificial neural network" first appeared in articles in the 1950s, supported by mathematical research inspired by the firing of neurons in the human brain.[83] Progressing from early mathematical models, computer scientists developed fields of study around machine learning and continue to improve their techniques. The combination of greater access to large networked databases and exponential advances in computing power, particularly graphic processing units, allowed for theoretical machine learning techniques, like deep learning, to become a

[82] Matt Leonard, "DARPA Looks to Control Drone Swarms with VR," *Defense Systems*, March 26, 2018, https://defensesystems.com/articles/2018/03/28/darpa-offset-drone-architecture.aspx.

[83] Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review* 65, no. 6 (1958): 386, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.3398&rep=rep1&type=pdf.

reality.[84] In 2015, computer scientist made substantial strides in machine learning by achieving human-level performance in an algorithm trained to play a wide variety of video games, demonstrating the technology's potential to rapidly find game-specific solutions.[85]

The concept of *machine learning* is a subset, though often cited synonymously, to the broader concept of AI. Machine learning describes a method to design a software algorithm where the person does not directly code the resulting decision-making logic.[86] Instead, the person specifies the parameter for inputs, outputs, and metrics for testing the results, and allows the machine to "learn" patterns through comparisons of training data. After sufficient iterations, the algorithm learns an optimal way to categorize or predict the best output given a new (untrained) input. What makes this technique powerful is that modern computers are able to iterate through millions of distinct algorithm configurations in short timespans.[87] Given proper parameters, and data for training, the result of using machine learning will be an algorithm that can reliably accomplish a specific task.[88]

Machine learning has three main techniques for training algorithms: supervised, unsupervised, and reinforcement learning.[89] Supervised learning occurs when the solution has a known answer for a particular input. The software can work backward from known output-to-input to determine the proper weights and biases to assign to the

[84] Tom Bianco, "GPU Acceleration Advancing the Evolution of Fast and Big Data," *Datanami*, July 24, 2017, https://www.datanami.com/2017/07/24/gpu-acceleration-advancing-evolution-fast-big-data/; Roger Parloff, "Why Deep Learning is Suddenly Changing Your Life," *Fortune*, September 28, 2017, http://fortune.com/ai-artificial-intelligence-deep-machine-learning/.

[85] Volodymyr Mnih, et al., "Human-Level Control through Deep Reinforcement Learning," *Nature* 518, no. 7540 (2015): 529, https://www.nature.com/articles/nature14236.

[86] Jaime Carbonell, Ryszard Michalski, and Tom Mitchell, "An Overview of Machine Learning," In *Machine learning: An Artificial Intelligence Approach*, eds. Ryszard Michalski, et al. (Berlin, Germany: Springer, 2013), 3–23.

[87] Richard Sutton and Andrew Barto, *Reinforcement Learning: An Introduction* (Cambridge, MA: MIT Press, 2014), 106, https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf.

[88] Mary Beth O'Leary, "Revolutionizing Everyday Products with Artificial Intelligence," *MIT News*, June 1, 2018, http://news.mit.edu/2018/revolutionizing-everyday-products-with-artificial-intelligence-mit-meche-0601.

[89] Carbonell, Michalski, and Mitchell, "An Overview of Machine Learning," 2013.

logic to get the desired results. Unsupervised learning is when there is no set result, but lots of data, and the programmer is relying on the software to suggest patterns in the datasets that may highlight hidden internal relationships. Reinforcement learning is teaching an agent how to solve a task in a simulated environment through trial-and-error and with rewards and punishments. While all machine learning techniques are useful in certain contexts, the emergence of reinforcement learning and its ability to train in an environment with little or no existing data, holds particular promise for training drone swarms how to operate.

The novel approach that Google used in 2017 with its AlphaGo Zero project was applying the machine learning technique of reinforcement learning to a data sparse game environment to solve a problem of unprecedented complexity.[90] In other words, the machine learning algorithm did not leverage an existing database of game solutions or previous information. Thus, with only the rules of the game, AlphaGo Zero experimented through self-play to accumulate generations worth of experience in the span of days and discover champion-level strategies. Similarly, applying machine learning techniques in data sparse environments related to mission-specific military wargames has the potential to change the character of war, specifically how future militaries will train to fight.

E. A WAY OF SWARM: STRENGTHS AND CRITIQUES

The combination of the existing DoD-sponsored frameworks that use wargames to simulate agent behaviors paired with machine learning techniques can produce a new framework that is rapid, flexible, and adaptive. This proposed Way of Swarm addresses the weakness of current drone training frameworks by incorporating the rising potential of machine learning and the lessons learned from the study of war. First, the combination of narrow AI with wargaming is rapid in execution and insulates the framework against the instability of crowd sourcing. The dependency on sustaining a crowd is replaced by

[90] Silver et al., “Mastering the Game of Go Without Human Knowledge,” 2017.

the persistent availability of AI and cloud computing.[91] Second, both the narrow AI and wargame are flexible and adaptive to changing assumptions, such as new hardware or environmental conditions. For instance, if the rules of the game of Go for some reason changed tomorrow (like adding a wall in the middle of the board), Google's framework could rapidly be run in days to again master the updated game. In a changing world of warfare, this rapid response to change is critical in design. Third, this framework applies top-down tactics development, but it does so by allowing the narrow AI to solve local drone interaction rules that optimize a global mission objective.[92] Hence, the framework preserves the principles of mission command and decentralized execution.

Despite strengths, there are also anticipated critiques for the proposed framework. First, producing and combining both high fidelity wargames and cutting-edge machine learning algorithms requires the combination of both the private sector's knowledge in AI and military subject matter experts. Unfortunately, friction around these issues has already occurred with nearly four-thousand Google employees demanding an end to their company's partnership with the defense department over Project Maven.[93] The project aimed to leverage narrow AI to reduce the human workload required to process, exploit, and disseminate collected intelligence, surveillance, and reconnaissance data. Therefore, to continue AI integration to enhance military projects, resolving these partnerships is a national imperative.

A second expected criticism is the risk of the "black box" phenomenon[94] of the swarm tactics produced by machine learning solutions. This phenomenon occurs when there is no rational explanation for the decision an algorithm makes due to the inherent

[91] "Cloud Services Support," Defense Information Systems Agency, accessed October 10, 2018, https://www.disa.mil/Computing/Cloud-Services/Cloud-Support.

[92] Istvan Fehervari and Wilfried Elmenreich, "Evolving Neural Network Controllers for a Team of Self-Organizing Robots," *Journal of Robotics*, vol. 2010, March 25, 2010, https://www.hindawi.com/journals/jr/2010/841286/abs/.

[93] Daisuke Wakabayashi and Scott Shane, "Google Will Not Renew Pentagon Contract That Upset Employees," *New York Times*, June 1, 2018, https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

[94] Will Knight, "The Dark Secret at the Heart of AI," *MIT Technology Review*, April 11, 2017, https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.

complexity of the machine learning technique used. The black box dilemma forces a balance between how much human bias to include in the swarm algorithm of the drones, and how much corresponding understanding is needed in the final algorithm. Ultimately, there is a balance in how much risk to assume. The trade-off is between low human bias and more decision-making flexibility, or high human bias that offers greater insight into why an algorithm made a particular decision, which can increase trust and confidence.

Finally, some critics of computer-based wargames contend that the assumptions inherent in computer-based models will produce an unacceptable gap between theory and reality.[95] According to their claim, this is due to the inability to capture human motivations like desire, commitment, passion, or will in simulation.[96] Although these critiques hold merit, it also depends on the intended use of the wargame. Typically, the more specific the computer-based wargame (less generalizable), the more the model is representative. Proponents also counter that when researchers validate wargames with additional methods, such as live experimentation, their utility for prediction is stronger.[97] Therefore, tailoring the wargame to a specific mission, with a defined set of assumption, rather than a wide range of tasks limits the expected swarm behavior and offsets this weakness. Additionally, validating the swarm algorithms produced by this framework with tests in the field adds another way to mitigate gaps between theory and reality.

[95] Roberto Leombruni and Matteo Richiardi, “Why are Economists Skeptical about Agent-Based Simulations?” *Physica A: Statistical Mechanics and its Applications*, vol. 335 (2005): 103–109, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.1155&rep=rep1&type=pdf

[96] Gabriel Almond and Stephen Genco, “Clouds, Clocks, and the Study of Politics,” *World Politics* 29, no. 4 (1977): 489–522, https://www.jstor.org/stable/2010037?seq=1.

[97] Manzo Gianluca, “Potentials and Limits of Simulation Multi-Agents: An Introduction,” *Revue Francaise de* Sociologie 55, no. 4 (2014): 653–688, https://www.cairn-int.info/article-E_RFS_554_0653--the-potential-and-limitations-of.htm.

III. MODEL CONSTRUCTION

SHALL WE PLAY A GAME?

—Joshua, *WarGames,* 1983[98]

The model constructed for this thesis is a custom-built swarming wargame that is playable by a machine learning technique, known as a genetic algorithm, and is scalable for rapid training through a cloud-computing service. The wargame incorperates lessons learned from autonomous systems research, historic training of decentralized forces, and recent advancements in machine learning, and combines them into an architecture that focuses on producing emergent swarm behaviors. Although the custom wargame is not a high-fidelity or fully-featured implementation of the final Way of Swarm framework, the model does provide a proof-of-concept that operates at an appropriately high-level of abstraction to serve as a thesis research tool. The primary purpose of this tool is to generate quantitative data to explore the following question: "How can wargames and machine learning be combined to train a decentralized swarm of autonomous systems, thereby enhancing the human-autonomy team?"

The following chapter is presented in four sections. First, the chapter discusses why the wargame was custom designed and why a genetic algorithm was selected for the model, as opposed to an existing high-fidelity gaming environment or a more advanced machine learning technique. Second, the chapter summarizes the design specifications for the wargame, to include the rules of the game, the layout of the gameboard, the assumed characteristics of the game agents, and the decision-making algorithm run by the agents. Third, the chapter outlines the genetic algorithm that accesses the wargame to help train key agent decision-making parameters for the game agents to generate swarm behaviors. Finally, the chapter examines the webserver that connects the wargame and the genetic algorithm, emphasizing how access to a scalable cloud-computing service incorperated in the model was essential to conduct millions of simulations for thesis research.

[98] *WarGames*, directed by John Badham (Los Angeles, CA: United Artists, 1983).

A. WHY BUILD A CUSTOM MODEL?

The literature review indicated that elements of decentralized mission command, rapid iterations through self-play, and that scalability to generate results in a data sparse environment are important when constructing a useful swarming training model. Initial efforts to find an existing industry standard wargame that was suited for pairing with machine learning, and also included these design elements, was unsuccessful.[99] The closest game identified to meet the research requirements was Swarm Commander,[100] produced by a coding team at Naval Postgraduate School—a leading research facility for swarming autonomous systems.[101] Swarm Commander allows the game user to control a collection of simulated drones by building logical scripts and via point-and-click user commands. The user moves swarms around a map and assigns them a pre-planned script that they execute based on a playbook that the user builds prior to playing the game.

Although the Swarm Commander game is a constructive tool to explore swarming concepts, it lacks three important elements for this thesis research. First, the design of the game focuses on a central controller (i.e., the user) to send commands to each drone swarm to facilitate game play. In theory, although the decision-making algorithm of a centralized machine controller could replace the human with a game modification, this would negate the ability to research benefits associated with training decentralized drone agents. Second, Swarm Commander runs in a real-time environment. The real-time display is ideal for allowing a human user to interact and update new commands to the swarms during gameplay, but this design reduces the capacity for a machine learning technique to rapidly simulate through thousands of games per minute to test different

[99] Nathan Padgett, "Defensive Swarm: An Agent-based Modeling Analysis," (master's thesis, Naval Postgraduate School, 2017), https://calhoun.nps.edu/handle/10945/56777; "MASON," George Mason University's Evolutionary Computation Laboratory, accessed November 12, 2017, https://cs.gmu.edu/~eclab/projects/mason/; "OFFensive Swarm-Enabled Tactics (OFFSET)," DARPA, accessed September 20, 2018, https://www.darpa.mil/work-with-us/offensive-swarm-enabled-tactics.

[100] "Swarm Commander Tactics," Naval Postgraduate School MOVES, accessed December 4, 2017, https://gitlab.nps.edu/moves/swarm-commander-tactics/tree/master.

[101] Matthew Schehl, "JIFX Continues to Help DoD, Academia Explore Limits of New Technology," August 9, 2018, https://www.dvidshub.net/news/295041/jifx-continues-help-dod-academia-explore-limits-new-technology.

tactics. Third, the game does not include the support for distributing and coordinated simulations across multiple computers (or multiple processors on the same computer) to scale the amount of data captured simultaneously. Therefore, without finding an existing wargame that met all the research requirements, the thesis team built a custom wargame.

The thesis team also used a customized adaptation of an existing machine learning technique for the research model. Google's AlphaGo Zero project used deep artificial neural networks and a deep reinforcement learning technique to defeat world champions in Go, a centrally controlled strategy game (i.e., a single player decides where to place all the pieces for each team).[102] Theoretically, the team could apply Google's deep learning techniques to a swarming wargame to develop decentralized tactics. However, the deep learning techniques still require additional research to determine how to modify the neural network parameters to train a decentralized group of agents that can coordinate for a collective objective. In other words, the goal for training a decentralized swarm is to train the agents themselves (i.e., each Go piece decides where to move or place itself to benefit the team), rather than to train a single, central player. Additionally, since artificial neural networks can generate a "black-box" solution to problems, neural network techniques can reduce trust and confidence in final solutions.[103] For these reasons, although artificial networks and deep learning shows promise for tackling decentralized swarm training in the future,[104] the model for this thesis uses an adapted machine learning technique known as genetic algorithms. Genetic algorithms are similar to artificial neural networks in their evolution-inspired approach to solving optimization problems.[105] Moreover, genetic algorithms are easier for researchers to explain in their execution and were more manageable to implement into the proof-of-concept wargame.

[102] Silver et al., "Mastering the Game of Go Without Human Knowledge," 2017.

[103] Knight, "The Dark Secret at the Heart of AI," 2017.

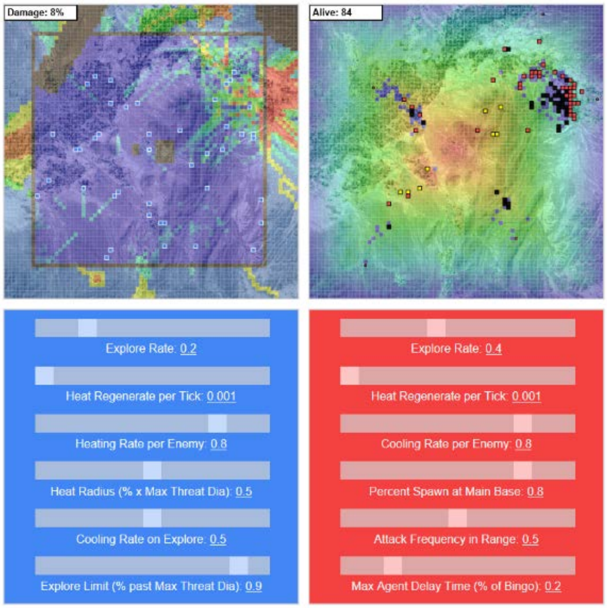
[104] James Vincent, "AI Bots Trained for 180 Years a Day to Beat Humans at DOTA 2," *Verge*, June 25, 2018, https://www.theverge.com/2018/6/25/17492918/openai-dota-2-bot-ai-five-5v5-matches.

[105] Robert Marks and Hermann Schnabl. "Genetic Algorithms and Neural Networks: A Comparison Based on the Repeated Prisoners Dilemma," In *Computational Techniques for Modelling Learning in Economics* (Boston, MA: Springer, 1999): 197–219, http://www.agsm.edu.au/bobm/papers/jena.pdf.

B. WARGAME: SWARMING AGENT-BASED MODEL

The wargame created is a JavaScript agent-based model that simulates high-level decisions, actions, interactions, and resulting emergent behaviors for a swarm of friendly Blue Force (BLUFOR) aerial drones and a swarm of enemy Red Force (REDFOR) ground threats in a hypothetical base-defense scenario. The inputs to the wargame are six key parameters for BLUFOR agents and six key parameters for REDFOR agents that affect each agents' local decision-making priorities. After the user, or machine learning algorithm, selects the agents' key parameters, the game creates a stochastic scenario that is bound by the wargame assumptions; it then simulates five hours of a base defense mission in a fraction of a second. The output of the wargame is a score for much damage the base received and represents the effectiveness of the swarm searching algorithm.

The wargame user interface allows a player to visualize the game environment, manually select the six key decision-making parameters per team, and step through the resulting simulations at a slower pace (Figure 2). The images on the top reflect the teams' respective *heatmaps*, or their collectively communicated picture of the world, and the slider-bars on the bottom depict the key decision-making parameters selected per team.



Damage: 8%

Alive: 84

Explore Rate: 0.2

Heat Regenerate per Tick: 0.001

Heating Rate per Enemy: 0.8

Heat Radius (% x Max Threat Dia): 0.5

Cooling Rate on Explore: 0.5

Explore Limit (% past Max Threat Dia): 0.9

Explore Rate: 0.4

Heat Regenerate per Tick: 0.001

Cooling Rate per Enemy: 0.8

Percent Spawn at Main Base: 0.8

Attack Frequency in Range: 0.5

Max Agent Delay Time (% of Bingo): 0.2

Figure 2. Wargame Interface

The primary focus of the wargame is how BLUFOR drones can best communicate and position themselves on a gameboard to search and detect a hostile REDFOR. Since the focus of the research is on high-level swarm behaviors, the wargame does not model aspects of drone flight dynamics, terrain avoidance, or environmental hazards (e.g., wind, dust, darkness). Instead, the wargame models the decisions that a drone makes to determine its next localized *move direction*, or on which grid tile that agent should next position itself to best support the swarming mission. This move direction is a high-level command that enters the drone's software for a lower-level autopilot software function (not modeled) to then determine how to physically maneuver the drone to the proper adjacent tile. Importantly, key local decision-making parameters for each team of agents are accessible through a machine learning technique that can then use consecutive games to learn what parameters optimally produce the most effective swarming behavior.

1. Rules of the Game

The wargame is turn-based, meaning all agents conduct their turns in order, based on each agent's calculated *action times*. Action times depend on agents' speeds and time delays for performing different actions. If two agents' action times are the same for their next action, then the wargame allows simultaneous actions; this ensures impartiality, so that no agent always receives a first-mover or last-mover advantage during turns. A full turn is complete when all agents execute their simulated decisions and actions. Full turns take less than a second to execute in-game, but represent fifteen seconds in real time.

The starting locations for all agents are controlled at the beginning of the game. All of the BLUFOR generate at the base at the start of the game. The REDFOR generate from randomly positioned REDFOR starting locations around the edge of the map, and they generate across a range of starting times (controlled by a key decision-making parameter). REDFOR knows where the BLUFOR base is located from the start of the game, but they do not get information about where BLUFOR drones will be searching. Additionally, BLUFOR does not have perfect knowledge of REDFOR starting locations or their attacking directions. For a BLUFOR agent to detect a REDFOR agent, it must position itself to be flying above them (i.e., in the same tile) and pass a detection check. This check is stochastic and depends on several factors outlined further in Section 2.

The top-down mission design chosen for the wargame is a base defense scenario. The mission objective for BLUFOR is to search and detect REDFOR, maintain a steady-state posture around their base, and minimize the damage inflicted to the base (Figure 3). The drone swarm assists in providing reconnaissance around the base and helps identify potential threats for other (notional) friendly assets to appropriately remove from the gameboard. The mission objective for REDFOR is to get within shooting distance of the BLUFOR base (within 15 game tiles) and attack the base to inflict damage.

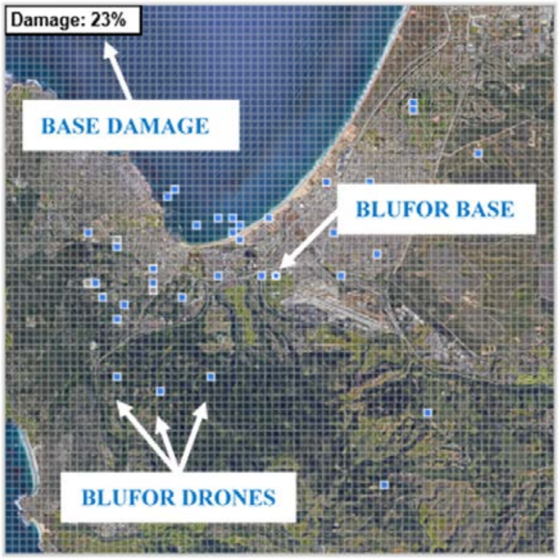


Figure 3. Base Defense Mission Overview[106]

Each swarm earns a score at the completion of each game based on how much damage the base sustains. The base sustains damage when a REDFOR agent is within shooting range of the base, chooses to shoot the base (probabilistic), and successfully hits the base (probabilistic). Each successful hit does a fixed 1% damage to the base. The base

[106] “Monterey Bay Map,” Google Maps, accessed January 20, 2018, https://www.google.com/maps/place/Monterey+Bay/@36.5870637,-121.8984165,29641m/data=!3m1!1e3!4m5!3m4!1s0x808e0ccfc5859dfd:0x124654a608855d43!8m2!3d36.8007413!4d-121.947311.

can never reduce its damage through repairs or regeneration. As such, the total base damage is a proxy variable to reflect how long it takes for BLUFOR agents to efficiently search and detect all the REDFOR agents on the gameboard that are an imminent threat.

2. Gameboard

The gameboard is a two-dimensional square grid that measures 7.6 by 7.6 miles. The board divides into 76 by 76 tiles (Figure 4). The satellite imagery in the background is a notional operating base, selected to provide the user a sense of scale. The gameboard does not contain terrain heights or objects (e.g., towers, buildings, trees) that might affect maneuverability of either force. Hence, the agents always have the option to move in any direction (including diagonals) or remain in their current tile. There are also no limits on how many BLUFOR and REDFOR can occupy the same tile location at the same time.

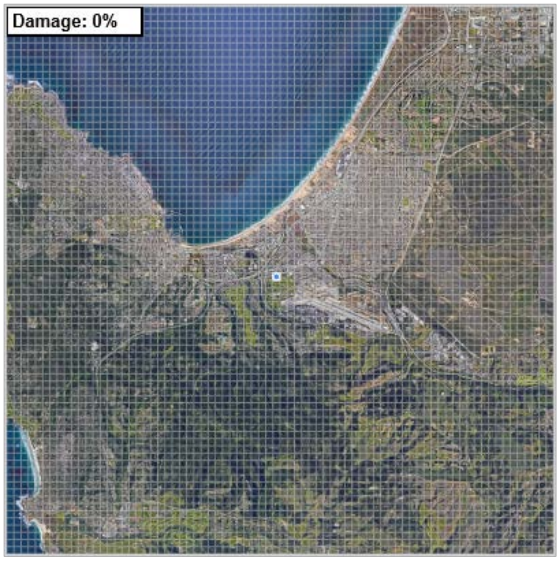


Figure 4. Example Gameboard

Each gameboard tile has an assigned *detection complexity score* that correlates to the probability that a drone would be able to find, fix, and track a ground target in that tile (e.g., an urban or forested area is scored as a higher complexity score than open terrain). The benefits of including a complexity score is that it makes the gameboard asymmetric, due to underlying map features, enhances the realism of the scenario, and makes it more challenging for BLUFOR to optimize on a single best solution for a search and detection tactics across the entire map. The detection complexity scores are rated as low, medium, or high, and are assigned manually based on a visual determination of the tile's terrain composition. The combination of the tile's detection complexity score, and whether or not the REDFOR is actively shooting at the BLUFOR base, generate a detection check scale for how likely a REDFOR will be detected by a BLUFOR flying above them.

Additionally, each gameboard tile has an *intelligence priority score* that is known by each BLUFOR agent at the beginning of the wargame. The intelligence priority score corresponds to a known location where intelligence analysts predict a higher likelihood of detecting REDFOR (Figure 5). BLUFOR agents access the intelligence priorities score in their local search regions to directly influence their individual and collective search patterns. Since intelligence assessments are not always perfect, the intelligence priority score does not correlate precisely with where REDFOR attacks will originate. However, there is a slighltly higher probability of REDFOR starting their attacks from the higher intelligence priority score locations (i.e., "hotter" tiles). The scores system also allows the user to specificy no-fly zones (black tiles) for sensitive regions around the map, such as local runways, and to set the search limit boundaries for the BLUFOR agents.

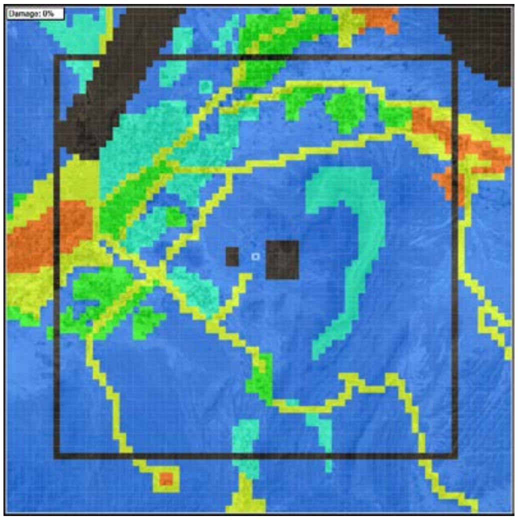


Figure 5. Intelligence Priority Scores

3. Agent Characteristics

Each REDFOR agent is a small enemy force that is demonstrating some form of hostile intent, or is conducting a hostile act, toward the BLUFOR base. They are best visualized as a mortar or sniper team. The REDFOR agents spawn from five locations around the edge of the map that correlate (weakly) to the intelligence priority score. REDFOR is able to detect BLUFOR agents one tile away (0.1 miles), which they can use to adjust the movement decisions for them and the rest of their team. All REDFOR agents have infinite ammunition. They have the choice to attack the base anytime they are in range, however shooting delays their time for taking their next action and it also increases their probability of detection. The chance of a REDFOR agent to successfully hit the base with an attack depends on their distance to the base. All successful hits on the base do 1% of damage. REDFOR cannot attack BLUFOR drones.

Each BLUFOR agent is a single drone assigned to help defend the BLUFOR base. They are best visualized as a small and inexpensive quadcopter with an autopilot system, video camera, and communications hardware (Figure 6). All BLUFOR agents start at the base in the center of the gameboard and launch over time to establish a steady state launch and recover cycle. Each drone has a finite battery life and will automatically fly a profile back to the base to replace its battery when it reaches the limit of its battery. A drone battery replacement at the base takes six minutes before the drone can relaunch.

Figure 6. BLUFOR Agent Representation[107]

[107] Exact Image from https://www.dji.com/phantom-4/info#specs.

The majority of the baseline assumptions for the characteristics of aerial drones and the capabilities of hostile ground threats to a base are based on the assumptions used in previous research.[108] The main agent characteristics are summarized in Table 1.

Table 1. Summary of Agent Characteristics

Team	Agent Characteristic	Value
BLUFOR	Maximum Speed	30 miles-per-hour
	Camera Search Rate	3.5 square-feet-per-second
	Battery Life	2 hours
	Battery Replacement Time	6 minutes
REDFOR	Maximum Speed	6 to 12 miles-per-hour
	Maximum Attack Range	1.5 miles (e.g., mortars)

The game also assumes that each drone flies with a proficient autopilot, avoids colliding with other drones, launches and lands by itself, communicates with all other drones, and uses onboard sensors to identify, fix, and track objects in the environment. Another key assumption, continued from previous research, is that BLUFOR are capable of distinguishing key features of REDFOR agents (e.g., weapons, military vehicles, etc.). Since the BLUFOR agents can rapidly make a determination of what activity appears hostile, there are no neutral agents represented in the wargame, as those (notional) neutral agents are already screened out. Although this is a significant assumption, the enhanced computer vision programs, like Project Maven, have shown this may soon be a reality.[109] Finally, data transmission between the agents is assumed to be short-range and redundant, and therefore inter-agent communication is never dropped, jammed, or incomplete.

[108] Padgett, “Defensive Swarms,” 2017.

[109] Pellerin, “Project Maven,” 2017.

Despite similarities in agent characteristics from previous research, the proposed model has two important differences in its wargame design. First, it does not rely on any centralized controlling authorities for executing commands. This enables the model to explore the benefits of a decentralized and self-organizing system of drones. Second, an emphasis in the scenario is that BLUFOR drones are strictly in search and detect roles. The drones pass any potential threats to (notional) friendly assets and the REDFOR agent is thereby removed from the board. The scenario does not intend, nor use, the drones to engage in any lethal strikes against the enemy. This is a deliberate design decision to illustrate that, without crossing ethical barriers, there are acceptable reasons for training autonomous systems, such as search missions, that can directly support the warfighter.

All game agents communicate to their teams using a decentralized technique for passing information that is similar to pheromone communication methods used by ants.[110] This pheromone method enables the agents to make local decisions about whether to add or remove pheromones, and about whether to follow or ignore pheromones, that results in a global swarm behavior as the teams disperse or converge around the gameboard. Each agent stores an internal pheromone map, visualized as a heatmap, of where the higher and lower tile regions of pheromones are located (Figure 7). The game represents each teams' heatmap as an overlay, where a tile that has more pheromones (higher priority) is darker red, and a tile that has fewer pheromones (lower priority) is darker blue.

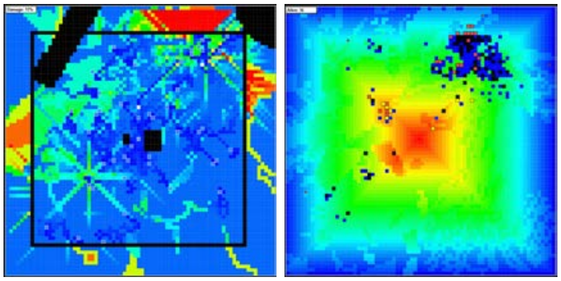


Figure 7. Agent Pheromone Communication Method (Heatmaps)

[110] J. L. Deneubourg et al., "The Self-organizing Exploratory Pattern of the Argentine Ant," *Journal of Insect Behavior* 3, no. 2 (1990): 159–168, https://link.springer.com/article/10.1007/BF01417909.

Agents update their heatmap each turn based on the communication reports they receive from other agents in their collective. Any agent can communicate that they have added or subtracted pheromones (i.e., heat) to tiles they visit due to decisions they make and the values they calculate based on the key decision-making parameters for each team. Importantly, all agents can only reference their individual construction of the collective heatmap; there is no single central reference map, and, in theory, individual heatmaps can diverge during mission execution. However, this model assumes heatmap divergence is negligible since each drone agent would use a secured, timestamped, and redundant communication method to ensure its messages transmit reliably across the entire swarm.

4. Agent Decision-Making Algorithm

The decision-making algorithm for all game agents is based on Colonel Boyd's observe, orient, decide, and act process (i.e., OODA loop). Each agent, in-turn, conducts an *observe phase*, *orient phase*, and *decide phase* of their loop based on the current state of their locally perceived environment. After deciding on what action to take, any agent that can take at least one action will simultaneously conduct the *act phase* of the loop.

In the *observe phase*, each agent checks its internal status and its local external environment for changes. For BLUFOR, each agent first checks its battery state to help determine if it needs to return back to base. Then, all BLUFOR agents use their cameras and sensors to conduct a search of the single tile directly underneath where they are located. If there are REDFOR agents underneath, then each BLUFOR in that tile has a chance of detecting them. The probability of detection depends on the tile's detection complexity score and whether the REDFOR agent is actively shooting at the base. For REDFOR, each agent first checks to see if they are within shooting range of the base. Then, all REDFOR agents scan all tiles adjacent to their current tile, trying to detect (i.e., see or hear) any BLUFOR drones. Finally, all agents from both teams reference and store the heat values of all their adjacent tiles (including the tile where they are located) to prepare them to orient the priority options for their next move directions.

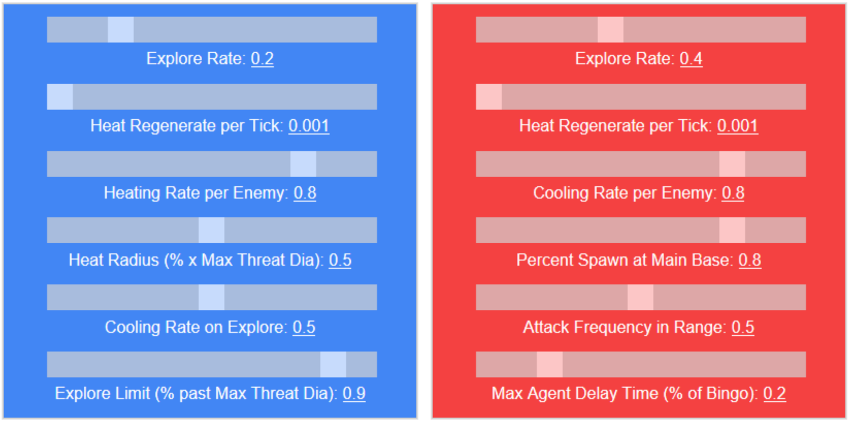
In the *orient phase*, each agent builds an internal heatmap, or operational picture, on which adjacent tiles are higher priority for their next move direction. For BLUFOR agents, regions are hotter based on the gameboard's intelligence priority scores, on a time-delay that incrementally adds heat to all gameboard tiles, and on drone communicated "starbursts" of heat (i.e., pheremone drops) that are centered around where REDFOR were discovered in previous turns. Regions are colder based on the chosen key decision-making parameter for heat removal rate that reduce the priorities of a tile for every agents that recently searched that tile. For REDFOR, regions with higher detection complexity scores and regions near the base are hotter on their heatmaps due to those being priority areas to shoot at the base. Areas get colder for REDFOR as they detect BLUFOR agents and alert others which tiles to avoid based on their observations. All combined, both teams observe the heat scores of all adjacent tiles, and then orient which tile among the nine options is the highest priority for their next move direction.

In the *decide phase*, each agent determines whether to move to a higher priority adjacent tile (as determined by building their heatmap), remain at the same tile, or move to a randomly selected adjacent tile. The key decision-making parameter that controls an agent's decide phase is the determined ratio of how often to explore versus exploit a given environmental scenario. To *explore* means to move in a random direction, whereas *exploit* refers to either staying at the current tile or moving to an adjacent tile based on which local tile has the observed highest priority (or heat) amount. Additionally, specific for BLUFOR agents, they will also decide whether or not to proceed directly back to base if they calculate their battery state is too low to continue searching. This decision to return to base to recharge will always override their decision to explore or exploit.

In the *act phase*, each agent conducts an engage, move, and communicate action. For BLUFOR, the engage action is detecting REDFOR with their cameras and sending video information back to other (notional) friendly agents to remove REDFOR from the gameboard. For the REDFOR, the engage action is kinetically shooting weapons at the base. Agents only shoot if they are in range of the base and they decide whether or not to risk shooting at the base. Additionally, only 10% of the shots REDFOR decides to take result in damage to the base. Shooting also increases the wait time until the next action, and increases probability of detection. For the move action, all agents will either wait in their current tile, or move to an adjacent tile, depending on the results of their decide phase. Finally, for the communicate action, all agents will send updates to their team that indicate which tiles they searched, what they detected, and whether or not to increase or decrease tile heat scores (and by how much) based on key decision-making parameters.

5. Key Decision-Making Parameters

The results of communicating and acting based on localized pieces of information among dozens or hundreds of drones executing hundreds of consecutive OODA loops forms the basis for a collective decentralized swarm. The decision-making process for the individual agents contains key parameters that affect how information is received, processed, acted upon, and further communicated to the rest of the swarm. Although a swarm could have hundreds of key parameters to manipulate, in demonstrating a proof-of-concept, this wargame extracted six key decision-making parameters per team that dictates their local interactions and, ultimately, impacts their global behaviors (Figure 8).



Explore Rate: 0.2

Heat Regenerate per Tick: 0.001

Heating Rate per Enemy: 0.8

Heat Radius (% x Max Threat Dia): 0.5

Cooling Rate on Explore: 0.5

Explore Limit (% past Max Threat Dia): 0.9

Explore Rate: 0.4

Heat Regenerate per Tick: 0.001

Cooling Rate per Enemy: 0.8

Percent Spawn at Main Base: 0.8

Attack Frequency in Range: 0.5

Max Agent Delay Time (% of Bingo): 0.2

Figure 8. Key Decision-Making Parameters

For visual representation, the key six parameters chosen for both teams' decision-making algorithms plot in a six-sided graph referred to as a swarm's *personality polygon* (Figure 9). The personality polygons have an axis for each key parameter that ranges from zero (center) to one (outside edge). Theoretically, any polygon that connects the six axes is a potential solution for setting the six parameters that would generate different emergent swarming behaviors. Given that each parameter is adjustable up to a precision of three decimal places, this puts the number of possible personality polygons at 1x10^{18}.

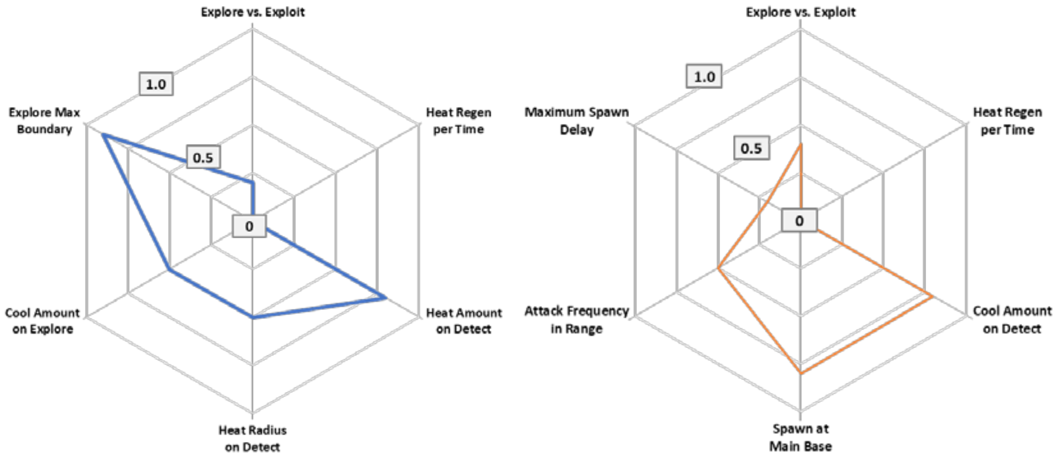


Figure 9. Example Parameter Polygons for BLUFOR and REDFOR

Based on the complex interdependent feedback mechanisms between individual drones and their emergent swarm behaviors, it is not easy to calculate or predict the best values to assign all the key parameters. Also, it is not feasible for a person to guess-and-check all combinations of possible polygons. Therefore, in order to train the swarms and find statistically better key parameters, the wargame was inentionally designed to allow a machine learning technique to systematically iterate through millions of wargames.

The first three parameters for BLUFOR and REDFOR dictate identical decision-making features. The *explore rate* parameters represent what percentage of the decisions an agent should make that exploits, or aligns with, the highest priorities of the collective (i.e., move in the direction of the hottest heatmap value), versus when to explore in a random direction. An agent that only exploits, would always position themselves where their collective heatmap was telling them was the highest priority location to move next. Yet, an agent that only explores, would always move randomly, and would negate all the efforts of the observing and orienting phases of their OODA loop. The *heat regeneration per tick* parameters represents how hot each tile should increase each turn to account for a decrease in certainty that an agent sufficiently and recently searched a tile. Too high or too low of a regeneration rate tends to result in a heatmap that is either fully hot or fully cold after searching, and thus the swarm is unable to communicate any other relative higher or lower priority search regions. The *heating/cooling rate per enemy* parameters represents how much an agent will communicate to heat/cool a tile to its collective swarm when detected. All heatmap tiles range in a heat score from zero (coldest) to one (hottest) and an agent cannot communicate them to be hotter or colder than these values.

Unique for BLUFOR, the *heat radius* parameter represents how large of a range (as a percentage of the assessed maximum REDFOR range) to communicate the heated starburst pattern. This parameter controls the distance the pheromone drop reaches, and influences the range that detection is likely to pull-in nearby drones to help investigate a region. If the radius is too large, the entire board tends to fill-up with excess pheromone. However, if the radius is too small, then the drones will not provide any pull to the rest of their swarm to direct them to leave their local exploring regions to support a globally hotter region with detected threats. The *cooling rate on explore* parameter represents how much heat an agent removes each turn as they search a tile. This factor reflects how confident an agent is that they successfully searched the tile, and how long until they believe another agent should return to search the tile. Finally, the *explore limit* parameter represents the furthest distance from the base that a swarm decides any agent should explore. The larger the limit, the more likely the swarm can detect REDFOR before they are within their maximum shooting range; however, a larger limit also means there is the more area for the swarm to search, thereby reducing the density of their search pattern.

Unique for REDFOR, the *percent spawn at main base* parameter represents the ratio of agents that will spawn at a single staging area. The higher the percentage, the more likely REDFOR will attack the base from a single direction as a consolidated mass. The lower the percentage, the more likely REDFOR will attack from multiple directions with a more dispersed force. The *attack frequency in range* parameter represents how often an agent will try to attack the base if they are within their maximum shooting range. Although REDFOR may desire taking shots more often, as it is the only means to damage the base, the action of shooting also delays their subsequent action phase for the agent and increases their likelihood of detection by BLUFOR agents. Finally, the *max agent delay time* parameter represents whether all agents should spawn and begin their attack maneuver toward the base at the same time, or whether they should spread out their attack over a few hours. Although attacking all at the same time increases the density of REDFOR for a single place and time, it can also be a weakness, since BLUFOR agents will tend to communicate for assistance when they detect a single REDFOR agent, thus drawing more drones to search in the dense regions of REDFOR.

C. MACHINE LEARNING: GENETIC ALGORITHM

The machine learning technique used to play the wargame is a modified version of a genetic algorithm. Genetic algorithms are a machine optimization technique that searches through an extensive range of possible solutions, referred to as *genomes*, using an evolutionary process that mimics natural selection.[111] For the swarming wargame, the genome is composed of the six key parameters (i.e., the personality polygons) for how the BLUFOR and REDFOR agents will make their decisions and interact at local levels. A batch of many different genomes is called a *population*. The genetic algorithm iterates through an evolutionary process that uses selection, crossover, mutation, and randomness to systematically test old populations and create new populations of genomes (Figure 10).

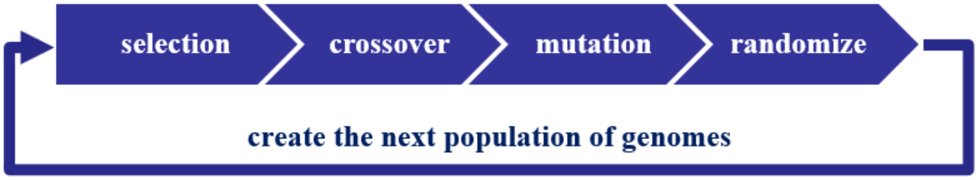


Figure 10. Genetic Algorithm Cycle[112]

Completing one cycle of the evolutionary process is called a *generation*. Every several generations, the top selected genomes are tested against an established baseline to

[111] Marks and Schnabl, “Genetic Algorithms and Neural Networks,” 1999.

[112] Modified from https://www.analyticsvidhya.com/blog/2017/07/introduction-to-genetic-algorithm/.

compare the progress of overall genome training. As the process continues over dozens, hundreds, or thousands of generations, although not guaranteed, the expected trend is that higher scoring genomes will continue to outperform the competition and succeed at being selected and re-populated as the most "fit" parameters for the training environment.

1. Hyperparameters

Machine learning techniques rely on user defined *hyperparameters*, or parameters that influence other parameters, in order to conduct iterations of training and testing on a given dataset.[113] For genetic algorithms, the primary hyperparameters deal with choosing a proper fitness function, selection criteria, crossover logic, and mutation rates. Although optimizing these values is, in itself, an area of active study,[114] this thesis research fixed hyperparameters based on effective values found in past research.[115]

A *fitness function* is the basis for how a genome is scored, and thereby influences which genomes are selected for further repopulation in subsequent generations of the evolutionary process. With the objective of this thesis to explore global swarm behaviors toward succeeding in a specific mission, the fitness function chosen was proportional to the total base damage sustained at the end of the wargame. The larger the base damage, the lower the fitness score for BLUFOR and the higher the fitness score for REDFOR. Moreover, since the wargame includes probabilistic events during the course of the game, an average of all fitness scores across a sample of multiple games played results in the overall fitness score assigned to each genome. These overall (average) fitness scores are used to determine which personality polygons are statistically better performing.

[113] Jesus Rodriguez, "Understanding Hyperparameters Optimization in Deep Learning Models: Concepts and Tools," *Medium*, August 8, 2018, https://towardsdatascience.com/understanding-hyperparameters-optimization-in-deep-learning-models-concepts-and-tools-357002a3338a.

[114] Rémi Bardenet, Mátyás Brendel, Balázs Kégl, and Michele Sebag, "Collaborative Hyperparameter Tuning," In *International Conference on Machine Learning* (2013): 199–207, http://proceedings.mlr.press/v28/bardenet13.pdf.

[115] Lee Jacobson, "Applying a Genetic Algorithm to the Traveling Salesman Problem," *Project Spot*, August 20, 2012, http://www.theprojectspot.com/tutorial-post/applying-a-genetic-algorithm-to-the-travelling-salesman-problem/5.

The hyperparameter chosen for *selection* is to pick the top 10% of the genomes by their overall fitness scores, mark them as *elites*, and to move those exact genomes into the next population batch without modification. This process of elitism helps to prevent taking steps backward in training, by ensuring there are some genomes at least as strong as previous generations continuing into the next population.[116] Elitism is also effective at validating previous genome results, since elites will experience different environmental conditions (e.g., enemy spawn locations) in the next generation of game iterations.

The hyperarameter chosen for *crossover* is to generate 20% of the next population from an "offspring" of sets of two randomly picked elite-genome "parents" from the previous generation. In the crossover process, one to five of the sequential values of the six key parameters are chosen from one parent genome, and the remaining values are chosen from a second parent genome to create a new genome. Crossover techniques are a method that systematically explores whether sequential pairs of values in a genome are synergistic and fitness enhancing.[117] In this wargame, since several of the parameters are interrelated (e.g., heat rate on enemy, heat radius on detect, and cooling rate on explore), crossover helps to search for genomes with effective pairings of these sequential values.

The hyperparameters chosen for *mutation* divide into two subgroups: *gross* and *fine mutation*. The gross mutation algorithm randomly picks elite genomes to generate 20% of the next population batch. The fine mutuation algorithm randomly picks elite genomes to generate 10% of the next population batch. Both algorithms have a 20% chance of mutating, or modifying, any of the six key parameters of the picked genome. The gross mutuation changes the key parameter up to a maximum of 10% from its previous value, whereas a fine mutation changes the parameter up to a maximum of 1%. Including a gross and fine mutation process helps a genetic algorithm explore whether moderate or small changes in an already successful genomes can further optimize to find

[116] Loris Serafino, *Between Theory and Practice: Guidelines for an Optimization Scheme with Genetic Algorithms—Part I* (Shenzhen, China: Kuang-Chi Institute of Advanced Technology, 2011): 16, https://arxiv.org/pdf/1112.4323.pdf.

[117] Ibid., 7.

a more precise solution, and to determine whether small disturbances in key values will have large emergent effects on the overall dynamic system (i.e., the butterfly effect).[118]

Finally, the hyperpareter used for *randomizing* the remaining 40% of genomes for the next population batch is through a pseudo-random algorithm that picks new values from zero to one for all key parameters. Randomization fills the remaining population back to its original batch size, so that the evolutionary cycle can perpetualy continue. In addition, with a large percentage of the next population batch being randomly generated, this technique helps to search for unexpected key parameter combinations, rather than merely exploiting the current best performing genomes. Randomization is used to prevent getting stuck with a set of personality polygons that may locally be producing the highest fitness scores, but are not the highest fitness score relative to the global solution space.

2. Initial Conditions

Research using genetic algorithms has found that the initial conditions used by the machine learning technique can have impacts on its ability to produce useful solutions for optimization problems.[119] Two important factors are the size of the starting population as compared to the size of the total solution space and the corresponding diversity in the distribution of the starting population. In other words, starting with too small of an initial population or too low of an initial distribution could result in a solution that gets stuck at a local versus a global optimal solution. Besides the quality of the fitness solutions found, the size and distribution of initial conditions also drives computational resources required to reach a solution. Evidence suggests that *quasi-random numbers,* or evenly-distributed numbers that still imitate randomness, provide a superior way to generate diversity within a population compared to pseudo-random numbers.[120] Therefore, the model uses quasi-

[118] James Richter, “On Mutation and Crossover in the Theory of Evolutionary Algorithms” (PhD diss. Montana State University, 2010), 75, https://www.cs.montana.edu/techreports/0910/Richter.pdf.

[119] Pedro Diaz-Gomez and Dean Hougen, “Initial Population for Genetic Algorithms: A Metric Approach,” In *GEM* (2017): 43–49, http://www.cameron.edu/~pdiaz-go/GAsPopMetric.pdf.

[120] Heikki Maaranen, Kaisa Miettinen, and Marko Makela, “Quasi-Random Initial Population for Genetic Algorithms,” *Computers & Mathematics with Applications 47, no. 12 (2004): 1885–1895, https*://core.ac.uk/download/pdf/82606936.pdf.

random numbers as initial conditions to reduce computational resources and to reduce the potential to get stuck in a locally optimal solution.

Applying the concept of quasi-random numbers, this thesis implements a Halton Generator[121] to produce an initial population of genomes for BLUFOR and REDFOR. Figure 11 provides a histogram for the initial population of BLUFOR used in generating the experimental data. It highlights the generator's effectiveness at distributing the initial values of the key parameters from zero to one across the six-dimensional search space.

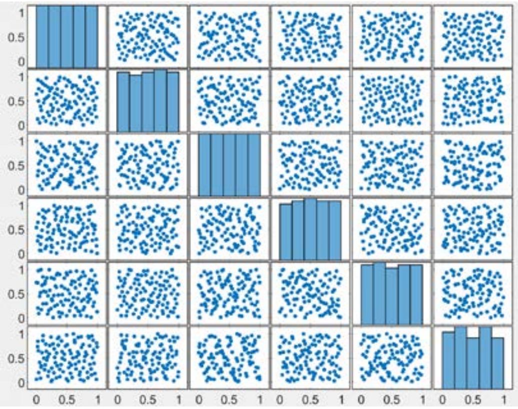


Figure 11. Histogram for BLUFOR Initial Conditions of 100 Genomes

D. SCALABILITY OF MODEL: CLOUD COMPUTING

Creating a custom model of a swarming wargame, and giving a genetic algorithm access to manipulate that wargame, was necessary to generate quantitative data to explore the thesis research question; however, the model was insufficient without the capacity to scale the model's architecture for rapid execution. Therefore, the research team designed the model to take advantage of a Naval Postgraduate School cloud-computing webservice located in the Coalition for Open-source Defense Analysis (CODA) Laboratory[122] that

[121] Heikki Maaranen, Kaisa Miettinen, and Antti Penttinen, "On Initial Populations of a Genetic Algorithm for Continuous Optimization Problems," *Journal of Global Optimization* 37, no. 3 (2007): 405, http://www.cs.uoi.gr/~lagaris/GRAD_GLOPT/projects/genetic_POPULATIONS.pdf.

[122] Barbara Honegger, "NPS 'Cloud Computing' Lab Up and Running," Naval Postgraduate School, February 8, 2010, https://web.nps.edu/About/News/NPS-Cloud-Computing-Lab-Up-and-Running-.html.

enabled centralized data collection and decentralized executions of millions of wargames. The result of gaining access to these powerful remote services was the ability to execute hundreds-of-thousands of complete wargame simulations in the span of a single day.

Aside from generating data, all data across a full generation of wargames played must be readily accessible for a genetic algorithm to work. The genetic algorithm needs the fitness scores for all genomes in a tested population (potentially thousands of games) before it can iterate through just one evolution of its selection, mutation, crossover, and randomization process to generate the following population. Based on this necessity, the model relies on a web database to store previous population results and to also post new population generations that are ready for the next batch of simulation. With this architecture, any computer with a web browser is able to communicate to that database and determine whether there is an untested genome from the new population batch that it should simulate. Then, after all simulations for that genome (i.e., a sample of wargames) is complete, the computer communicates the simulation results back to the database and the process repeats. Thus, the more computers, the faster the output of the framework.

Pairing the strength of a web-based database for dataset management, with an architecture that directly scales by adding more computing resources, the model rapidly expanded with the addition of cloud-computing access. The Naval Postgraduate School CODA Laboratory offered remote-access to a hosted cluster of 16 computer processors and 64 gigabytes of memory, which allowed 40 simultaneous instances of the wargame to be run for several straight days. Each instance completed an entire wargame (simulating five hours of mission execution) in approximately ten milliseconds. Access to cloud-computing scaled the training architecture to execute millions of games over the course of a few days. This capability replaced the need to build a human crowd or to rely on personal computers to manually execute wargames. Ultimately, the custom architecture and model design enabled the generation of sufficient data for quantitative research.

IV. METHOD AND ANALYSIS OF RESULTS

The experimental design for this thesis combines the proof-of-concept drone swarm wargame and the machine learning technique outlined in Chapter III to generate data for a quantitative analysis of the proposed Way of Swarm framework. To make this assessment, the machine learning technique self-played two million iterations of the wargame in a co-evolving format. This method permitted a statistical comparison of the performance of subsequent generations of the swarming algorithm against a reference baseline adversary. Ultimately, the experiment produced statistically significant results that address the thesis research question by demonstrating the ability to rapidly train a swarming algorithm that (1) accomplishes a specific mission objective and (2) reduces the workload of the human to operate a swarm. Additionally, the experiment highlights the rapid, agile, and flexible benefits of using machine learning techniques through co-evolving self-play. The results show that the framework can generate an effective swarm algorithm, in only a few training days, while also being able to adapt to different initial conditions and environmental assumptions.

The following chapter is presented in two sections. First, the chapter outlines the experimental design for testing the model. This section discusses the control and test variables, the requirement for generating a baseline, a test of the model's sensitivity to initial conditions, and the co-evolutionary process for training swarms. Second, the chapter presents the experimental results and data analysis. This section includes the optimized key decision-making parameters for swarming agents for the wargame, an in-depth interpretation of those parameters, and a statistical analysis of the model and the relative change in swarm algorithm effectiveness.

A. EXPERIMENTAL DESIGN

1. Control and Test Variables

To assess which values for the key decision-making parameters produced the best results for each mission, the experimental design isolated individual agent parameters (i.e., six for BLUFOR and six for REDFOR agents) as independent variables in the model. Additionally, the experiment either fixed or controlled for all other factors that impact the dependent variable (i.e., base damage) within a set range of assumptions. Table 2 outlines the controlled, independent, and dependent variables of the experiment.

Table 2. Control and Test Variables

Controlled	Rules of the Game (i.e., Mission, Probabilities, Base Damage)
	Game Board (i.e., Board Dimensions, Human Intelligence Priorities, Terrain)
	Agent Characteristics (i.e., Speed, Number, Sensor Range)
	Genetic Algorithm Hyper-parameters (i.e., Elite, Cross-over, Mutation)
Independent	Agent Decision-Making Algorithms (i.e., Specific Parameters - Six Each)
Dependent Variable	Fitness Score (Percent Base Damaged)

Leveraging existing military planning techniques, like the Military Decision Making Process (MDMP), the experimental design focused training on a specific mission set, environment, and scenario. The experiment also applied a concept used by military planners to frame the adversary's capabilities, known as *most dangerous course of action* (MDCOA).[123] This concept, which is similar to a prudential strategy in game theory, refers to finding a worst-case scenario by evaluating the adversary's strategic options and their assessed capabilities.[124] Using an intelligence assessment of the adversary, military planners develop the MDCOA from the strategic to tactical levels of war to develop friendly courses of action to maximize mission success. Likewise, this thesis applied this concept to help train BLUFOR by setting REDFOR to a (notional) assessed MDCOA. In

[123] Department of the Army, *Commander and Staff Organization*, FM 6-0 (Washington, DC: Department of the Army, 2017): 9-15, http://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN7501_FM_6-0_Incl_C2_FINAL_WEB.pdf.

[124] Philip D. Straffin, *Game Theory and Strategy* (Washington, DC: Mathematical Association of America, 1993): 70.

this case, the REDFOR MDCOA consists of training against 150 agents with 5 attack locations and the capabilities for REDFOR agents as listed in Chapter III, Section B.

Additionally, to prevent the machine learning technique from "overfitting" a BLUFOR solution based on human intelligence priorities or on fixed starting locations of the REDFOR, the experiment stochastically adjusts REDFOR starting locations within a fixed range for each game. For instance, the wargame requires that at least one of the five starting locations is within an area marked as a high-confidence threat location via the human intelligence assessment. The remaining four starting locations must start at least 1.5 miles away from the BLUFOR base, but otherwise they have no further restrictions. This helps to reduce the overfitting of BLUFOR key decision-making parameters based on "memorizing" a known starting location. Therefore, although the gameboard location and terrain map remained fixed, the experimental design used a controlled variation within the environment to generate a more robust and generalizable algorithm solution.

2. Generating a Baseline

To assess the performance of subsequent generations of BLUFOR as the training iterations progressed, the experiment required a stable reference point for a comparison. Simply using the BLUFOR fitness score (i.e., average base damage) for each game would not be an accurate representation of the performance of the algorithm effectiveness since REDFOR is also improving for each generation. Google solved this challenge in its co-evolving gaming framework by using the Elo rating system to assess the performance of its algorithm.[125] Elo is a self-correcting scoring system, often used to measure the relative skill level of players in games like Chess or Go, based on past performance.[126] However, since this technique weighted the points a winner received based on the Elo rating of the opponent (self-correcting), the best way for Google to increase the rating of its algorithm was to play opponents with higher Elo ratings. This thesis did not use the Elo rating

[125] Silver et al., "Mastering the Game of Go Without Human Knowledge," 2017.

[126] Paul Albers and Han Vries, "Elo-rating as a Tool in the Sequential Estimation of Dominance Strengths," *Animal Behavior* 61 (2001): 489, http://dx.doi.org/10.1006/anbe.2000.1571.

system, as each genome would have required an associated Elo rating through all training cycles, increasing the data management requirements and complicating the design.

Rather, applying the general concept of Elo ratings, the experimental design used a simplified progress-check against a known baseline. For example, instead of measuring a continuous swarm rating, the experiment used a trained group of elite REDFOR from the results of an initial run of the framework to serve as a baseline adversary. The elite REDFOR genomes represented a formidable set of opponents, embodying the concept of MDCOA due to its evolutionary stability, meaning these REDFOR personality polygons did not significantly change with additional training evolutions. This corresponds to their fitness scores not improving with further modifications to local interaction parameters. Figure 12 shows the evolution of personality polygons for the elite group of REDFOR genomes used as a baseline for future training. The shape of personality polygons after 1,040,000 games does not significantly change, compared to those after 540,000 games, supporting the claim that they had evolved to a reasonable MDCOA level of proficiency.

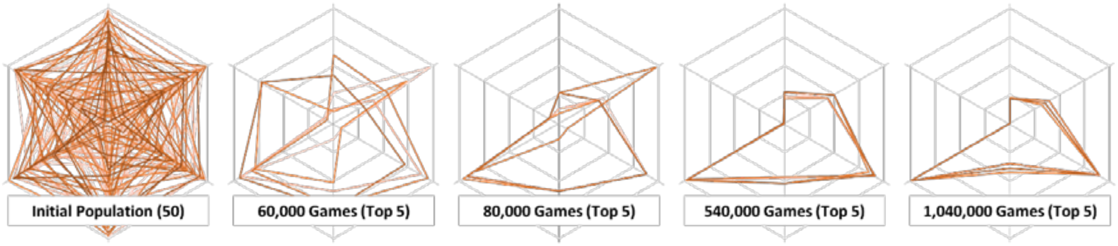


Figure 12. Generating a MDCOA Baseline for REDFOR

To expedite the development of a baseline adversary, the MDCOA creation matches limited the initial population for BLUFOR and REDFOR to 50 genomes, instead of 100, used in the final experiments. This allowed for generating the baseline in less time.

3. Sensitivity to Initial Conditions

Research in optimization suggests that the solutions found by genetic algorithms, as well as other machine learning techniques, potentially have a strong dependence on the initial conditions (i.e. initial population of genomes).[127] Therefore, to assess the severity of this effect on the model, this research performed a second iteration of the training framework, with a different set of initial conditions, to create a second MDCOA baseline for REDFOR. Figure 13 shows the personality polygons for REDFOR that resulted from using two different initial conditions. Although there are differences in the evolutionary process to reach the solutions, both iterations produced final polygons that display a high degree of similarity. This adds confidence that the solutions this custom wargame and machine learning technique produce are not highly sensitive to the initial conditions.

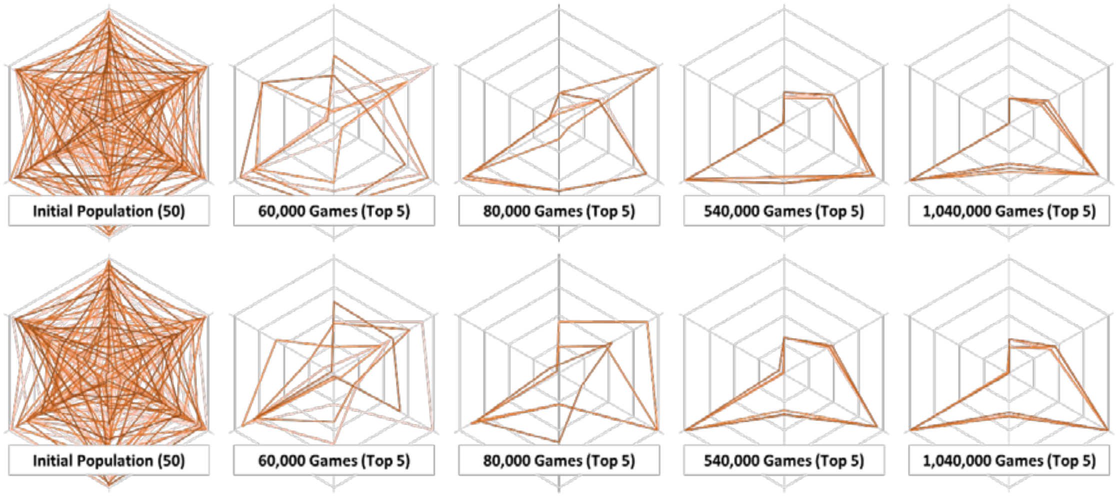


Figure 13. Sensitivity Analysis: Generating a MDCOA Baseline for REDFOR with Two Different Initial Conditions

Therefore, the final MDCOA baseline for REDFOR (Figure 14) combined results from the initial and sensitivity test into a single set of opponents for training progress-checks.

[127] Maaranen, Miettinen, and Makela, "Quasi-Random Initial Population," *2004*; Diaz-Gomez and Hougen, "Initial Population for Genetic Algorithms," 2017.

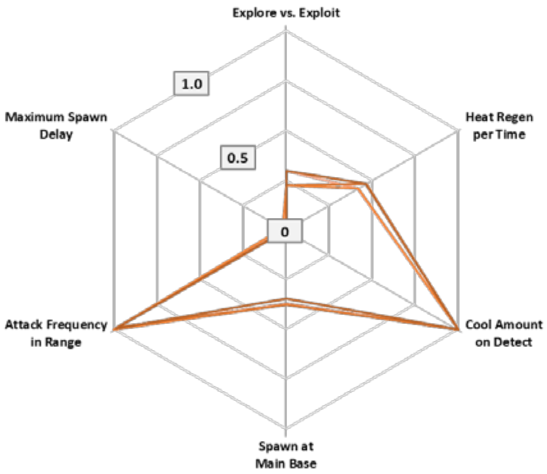


Figure 14. Personality Polygons of MDCOA Baseline for REDFOR

4. Sequential Co-Evolution

Applying the concept of self-play, adopted from Google's AlphaGo Zero[128], this research's experimental design also enabled both BLUFOR and REDFOR to learn as the training evolutions progressed. However, to best control the training in a deliberate manner, each force co-evolved in an alternating sequence. To begin, both REDFOR and BLUFOR initial populations played and then simultaneously evolved new populations. After these matches, the next population of BLUFOR genomes played against the elite REDFOR genomes from the first generation. Next, the roles reverse, and the population of REDFOR genomes played the elite BLUFOR genomes from the second generation, and so forth. Table 3, depicts the chosen co-evolving sequence that alternates generations of populations against an opponent's elite genomes.

Table 3. Co-Evolving Sequential Training Iterations

	1st Generation		2nd Generation		3rd Generation	
	Genome Population	Elite (top 10%)	Genome Population	Elite (top 10%)	Genome Population	Elite (top 10%)
BLUFOR	100	N/A	100 Evolved	10	10	N/A
REDFOR	100	10	10	N/A	100 Evolved	10

[128] Silver et al., "Mastering the Game of Go Without Human Knowledge," 2017.

The experimental design used the genome population sizes as listed in Table 4. A primary tradeoff in selecting the population size was balancing the desire to generate statistically significant findings against the ability to run a large quantity of wargames in a short time. Since the best solutions for both BLUFOR and REDFOR genomes are unknown prior to running the first generation, the experimental design started with 100 initial genomes for both forces. Subsequent generations applied the sequential co-evolution process, limiting one force to only the elite (top 10) genomes to compete against a new population (100) of the adversary’s evolved genomes. Additionally, every match between each distinct BLUFOR and REDFOR genome conducted twenty games. This allowed for an average fitness across the twenty sample games, which accounted for the stochastic nature of the wargame and the variation in REDFOR starting locations.

Table 4. Genome Population Sizes and Game Sample Sizes

Generation	Genome Size		Games Per Matchup	Total Runs	Cumulative Runs
	BLUFOR	REDFOR			
1	100	100	20	200,000	200,000
2	100	10	20	20,000	220,000
3	10	100	20	20,000	240,000
…	100/10	10/100	20	20,000	…
30	100	10	20	20,000	780,000

B. ANALYSIS OF EXPERIMENTAL DATA

To assess the potential of the framework to turn a collection of decentralized drones into a swarm, the experimental design used 780,000 iterations of the wargame, conducted over two days, evolving both BLUFOR and REDFOR over 30 generations. Every five generations, the elite BLUFOR genomes played the baseline REDFOR genomes to obtain a progress-check on the performance of the swarming algorithm. The training framework concluded when there were no longer significant changes in the fitness scores of subsequent baseline progress checks. Figure 15 depicts the fitness scores (i.e., average base damage) plotted for each progress check across the 780,000 games.

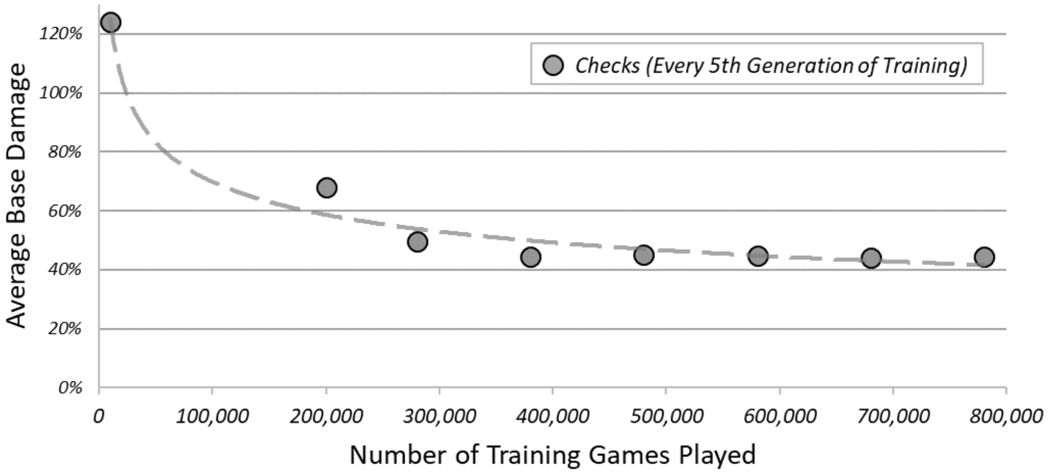


Figure 15. Average BLUFOR Base Damage as Progress Checked Against MDCOA Baseline for REDFOR

Figure 16 shows a visual progression of personality polygons for BLUFOR as the genetic algorithm processed through the solution space to find optimized parameters. The visualization starts with the initial population of pseudo-random genomes, then shows the first set of elite BLUFOR (top performers in the initial matchups), then the first set of their repopulated (or evolved) genomes, and so on. The expansion and contraction of the polygons is a product of the alternating co-evolutionary experimental design, and the hyperparameters selected for evolutionary elitism, crossover, and mutation in the genetic algorithm. Although this process can be computationally expensive, it also highlights the strength of a genetic algorithm to search a space methodically to find an optimal solution.

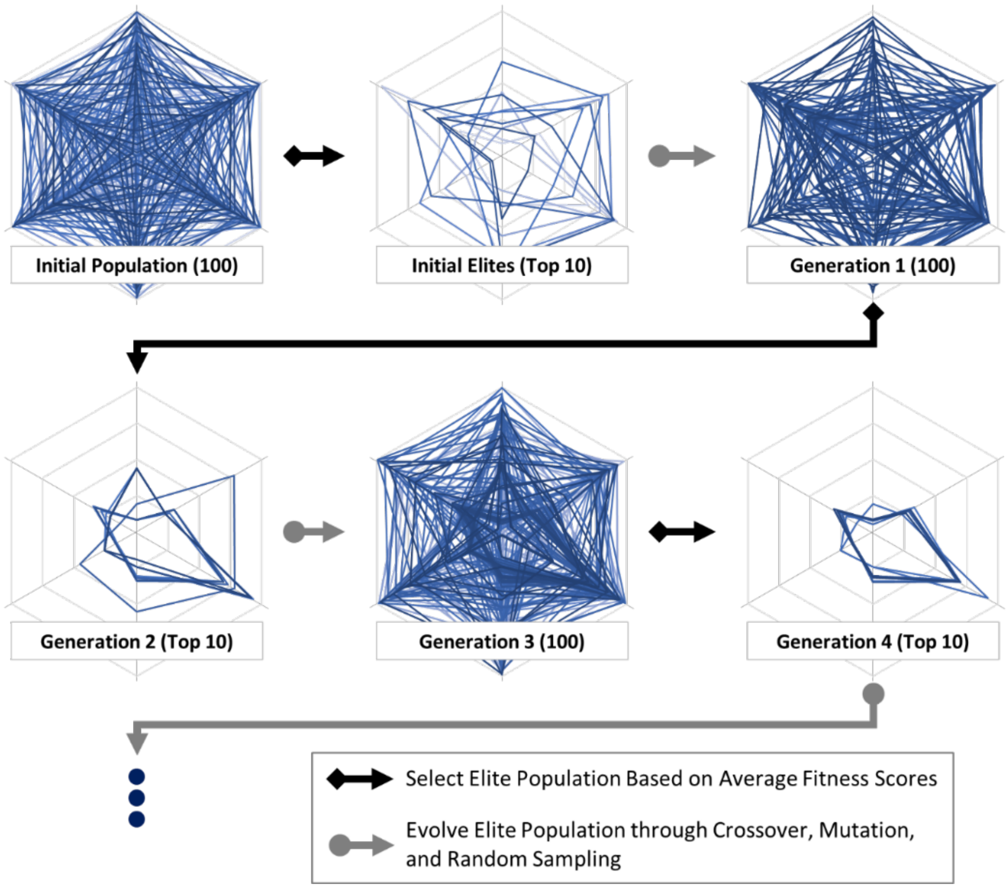

Initial Population (100)
Initial Elites (Top 10)
Generation 1 (100)
Generation 2 (Top 10)
Generation 3 (100)
Generation 4 (Top 10)
Select Elite Population Based on Average Fitness Scores
Evolve Elite Population through Crossover, Mutation, and Random Sampling

Figure 16. Genetic Algorithm Visualized through BLUFOR Evolution

1. Experiment Results

The result of the training framework was the output of optimized values for the best set of the six key decision-making parameters for both BLUFOR and REDFOR. These six decision-making parameters drove the local interaction rules of each individual agent, which translated into the accomplishment of a desired global behavior (e.g., for BLUFOR, minimizing average base damage). This section presents the results of the experiment by highlighting the evolutionary progression of the personality polygons that culminated in a final set of best-trained parameters for the mission.

Initially, BLUFOR and REDFOR started the experiment with 100 unique quasi-randomly generated genomes, shown in Figure 17, representing the initial populations.

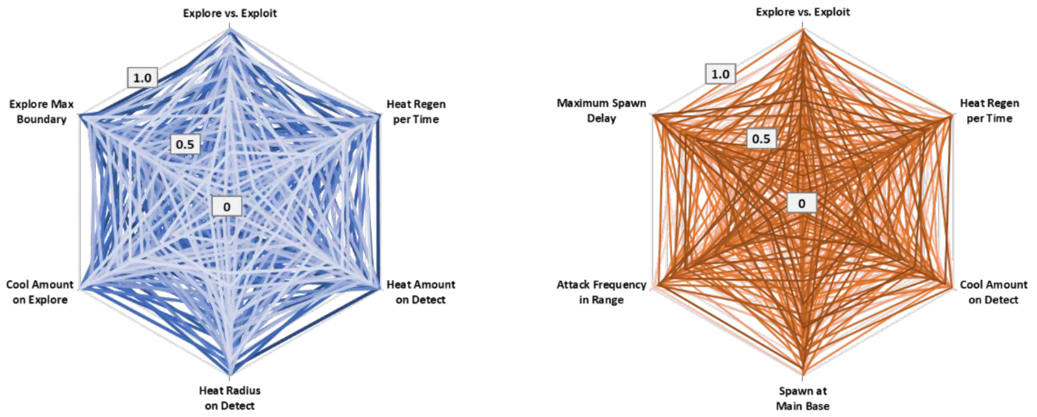


Figure 17. Initial Populations of 100 Quasi-random Genomes

From this initial set of distributed samples in the search space, the experimental design enabled each genome of BLUFOR to play 20 iterations against each genome of REDFOR, totaling 200,000 wargames. The fitness scores (i.e., average base damage) for each match of 20 games determined the elite (top 10) genomes for each force. Figure 18 depicts this initial generation of elite genomes per force. The variation of each set of elite genomes within the initial generation highlights how the genetic algorithm robustly worked through the solution space. From the initial elite populations, the experiment continued, performing the sequential co-evolution method for the remainder of the trial.

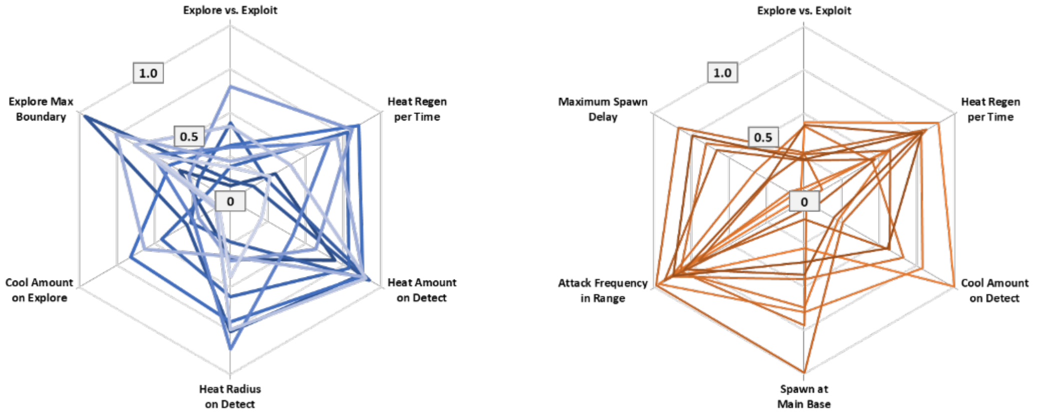


Figure 18. Elite Personality Polygons after 200,000 Games

Ultimately, after 780,000 wargames, or 30 generations of co-evolutionary training, the framework produced the best-trained genomes for both forces. Figure 19 displays the resultant elite personality polygons for both BLUFOR and REDFOR.

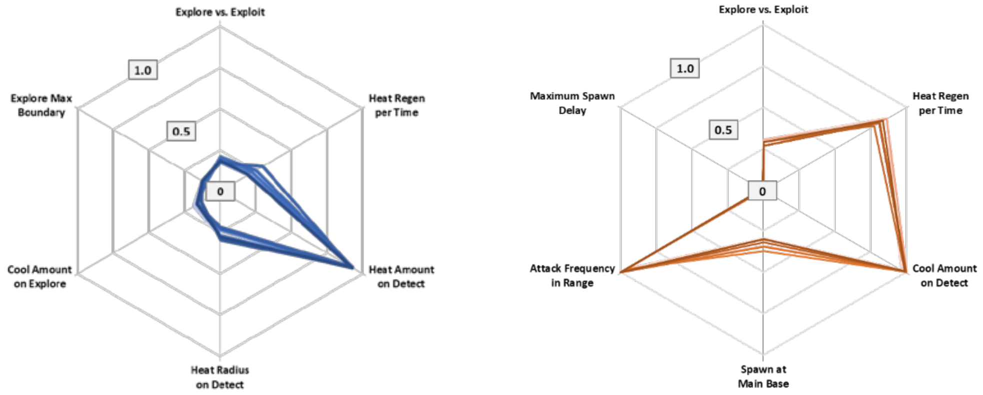


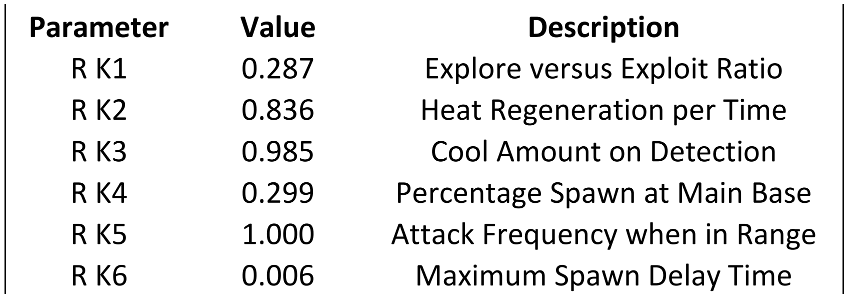
Figure 19. Elite Personality Polygons after 780,000 Games

Tables 5 and 6 provide the specific values for the key decision-making parameters for the best-trained BLUFOR and REDFOR genomes. These values are responsible for the local interaction decisions of each set of individual team agents, which translated into the best accomplishment of the desired base-defense mission by the swarm in the wargame.

Table 5. BLUFOR Key Decision-Making Parameters after 780,000 Games

Parameter	Value	Description
B K1	0.208	Explore versus Exploit Ratio
B K2	0.252	Heat Regeneration per Time
B K3	0.924	Heat Amount on Detection
B K4	0.231	Heat Radius on Detection
B K5	0.159	Cool Amount on Explored
B K6	0.105	Exploring Maximum Boundary

Table 6. REDFOR Key Decision-Making Parameters after 780,000 Games



Parameter	Value	Description
R K1	0.287	Explore versus Exploit Ratio
R K2	0.836	Heat Regeneration per Time
R K3	0.985	Cool Amount on Detection
R K4	0.299	Percentage Spawn at Main Base
R K5	1.000	Attack Frequency when in Range
R K6	0.006	Maximum Spawn Delay Time

2. Interpreting the Data

The proof-of-concept for this research used a high level of human bias in the development of the wargame, which also affords a high degree of insight into the results. In other words, since the game used key decision-making parameters that human programmers conceived, the resultant personality polygons are straightforward to interpret and justify. This increases trust in the resulting algorithm by providing the ability to explain how a trained swarm will likely behave in the operations, which is a valuable attribute when seeking to form an effective human-autonomy team. This also increases trust when human operators deploy a trained swarm by helping them understand the left and right limits, or constraints, of a trained decentralized swarm. Since the swarming algorithm trained for a specific mission, adversary, and environment, it is beneficial to examine the trained parameter values to help understand future behavior from within this narrow context. This section examines the values of the three parameters that are common to both BLUFOR and REDFOR, then analyzes the three unique to BLUFOR, and concludes by assessing the three parameters distinct to REDFOR.

First, this analysis examines the best-trained values within the shared parameters of BLUFOR and REDFOR genomes (i.e., first three parameters listed in Tables 5 and 6). The genetic algorithm nearly maximized the parameter for the *heating or cooling amount on detection* (depending on REDFOR or BLUFOR) for both forces (0.924 for BLUFOR and 0.985 for REDFOR). This maximization makes sense for both forces. For instance, for BLUFOR, finding REDFOR is the primary way to minimize average base damage, so the term must have a significant weight to drive the collective swarm behavior. This means that when a BLUFOR agent finds a REDFOR agent, it produces a strong signal to alert the other members in the swarm. Similarly, avoiding detection by BLUFOR is the primary way for REDFOR to maximize the number of agents available to attack the base. Therefore, when REDFOR agents detect BLUFOR agents, they produce a strong signal to warn the other members in the swarm. This concept is similar to the way ant foraging dynamics work through the use of pheromones.[129] The stronger the pheromone, the stronger the collective behavior will be, producing a cascading effect.

Next, the-*explore-versus-exploit* terms for both BLUFOR and REDFOR are very similar (0.208 for BLUFOR and 0.287 for REDFOR). This balance is in line with other research efforts in evaluative feedback, which emphasizes the need to allow an agent to explore, rather than just exploit, what it initially perceives as the best option to enable the discovery of a better solution.[130] The trained values for BLUFOR and REDOR support this finding, but they also highlight that the individual members of a swarm are not particularly efficient in and of themselves. However, although the high rate of exploration can, at times, produce overlap, the power of a swarm does not reside in creating highly efficient agents at the individual level, but effective behavior on the global scale. Therefore, the vales for which the genetic algorithm solved through the framework make sense and are consistent with previous research in evaluative feedback.

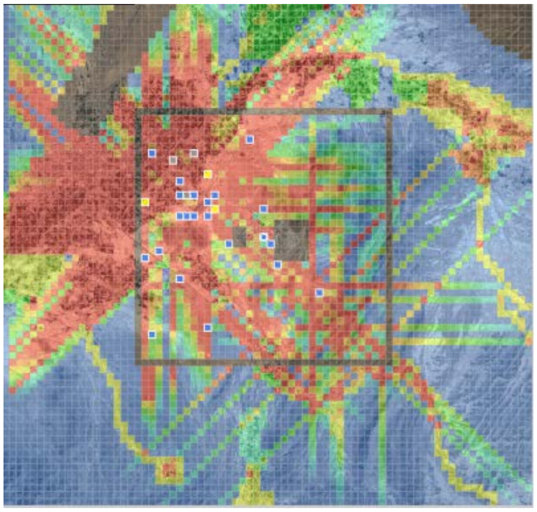
The last common parameter between BLUFOR and REDFOR is *heat regeneration*, which displays a significant difference in final values (0.252 for BLUFOR and 0.836 for REDFOR). However, both values make sense by examining them in the context of the wargame. First, the value selected for BLUFOR indicates that a lower rate for heat regeneration was beneficial. This makes sense since this parameter predominantly affected tiles with a preset intelligence priority (i.e., a location in the

[129] Erol Sahin, *Swarm Robotics: From Sources of Inspiration to Domains of Application*, Report Numbers METU-CENG-TR-2005-01 (Ankara, Turkey: Middle East Technical University, 2005), http://www.kovan.ceng.metu.edu.tr/pub/pdf/METU-CENG-TR-2005-01.pdf.

[130] Howard M. Schwartz, "Multi-agent Machine Learning: A Reinforcement Approach" (Hoboken, NJ: John Wiley & Sons, 2014), 52-56, https://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=1775207.

environment assessed as a likely location for REDFOR activity). However, the conditions of the game only required one out of five of the REDFOR starting locations to generate within these areas. Thus, it follows that the final BLUFOR genome would place a low value on this parameter, opting to search and react to REDFOR actions, rather than bias toward assessed threat locations. Conversely, REDFOR placed a high value on heat regeneration. This weighting is logical because BLUFOR drones travel at a much faster relative speed. This requires a dynamic operating picture to enable REDFOR to take advantage of fleeting opportunities in the relative positions of BLUFOR. For instance, if the heat regeneration had been lower, REDFOR would have perceived favorable moves as unfavorable (e.g., BLUFOR is no longer on the cooled tile). In other words, since the movement speed of BLUFOR drones was fast, the heat regeneration needed to be fast to accurately represent the environment.

Second, this analysis examines the specific values for the parameters unique to both the best-trained BLUFOR and REDFOR genomes. Starting with BLUFOR, there were three unique parameters. First, the value for the *heat radius on detection* (0.231) controlled how far an agent would transmit heat from a location where it located a REDFOR agent. If the value of this parameter became too large, the heat map quickly became saturated with heat (Figure 20). The saturation effect inhibits effective swarm behavior by making it difficult to discern the location of REDFOR. Therefore, the final trained value balanced the need to increase the spread of information against the need to prevent an over-saturation of heat.

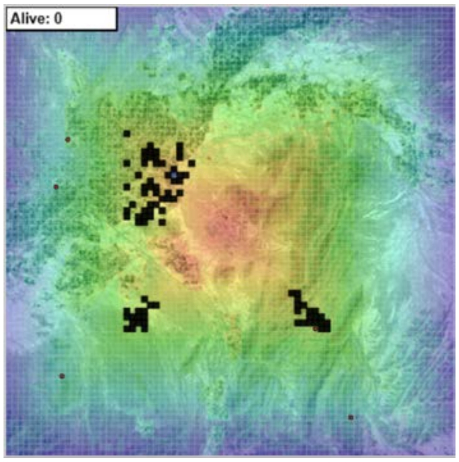


Increasing the heat radius oversaturates the heat map of individual drones by making nearly all the tiles appear as a high priority (red). This effectively increases random behavior vice effective swarm behavior.

Figure 20. Oversaturation of Heat Pheromone

Next, the low value for the *cooling rate on explored* parameter (0.159) indicates that it was beneficial to place low confidence in an individual agent's ability to search a tile (low cooling value). Consequently, this reduced how long it took for another agent to return to re-search the same location. Decreasing the revisit time between searches makes sense considering that the probability of detecting REDFOR agents is not perfect and that REDOR agents are constantly moving. Finally, the trained value for the *explore limit* (0.105) was the lowest value in the BLUFOR genome. This value indicated that it was beneficial for the BLUFOR agents to explore out to the maximum expected threat range of REDFOR, but not much farther. This is not surprising, because the sensor range for each BLUFOR was limited to one tile per time-step, thus the only way to maintain a sufficient search density of sensors was to heavily restrict the operating area for BLUFOR. It is logical to infer that increasing the number of BLUFOR drones or their sensors' range, for this given scenario, could increase the value of this parameter.

In contrast, REDFOR had three distinct parameters. First, the value for the *percent spawn at main base* (0.299) indicated that it was beneficial for REDFOR to mass, to a degree, but distributing its force across all five locations to produce a simultaneous and omnidirectional attack provided the greatest chance for success. This is not surprising, since defending against this type of attack, referred to as a saturation attack, often proves difficult for military planners.[131] Next, examining the *attack frequency in range* (1.000), a high value for this parameter was expected, as shooting is the only means to damage the BLUFOR base. However, a maximum value was not expected. The act of shooting slows REDFOR agents' subsequent action phase and increases their likelihood of detection in the wargame, but this did not appear to inhibit the decision to shoot as much as possible when entering effective range. It is possible this is because REDFOR learned that the density of BLUFOR agents diminishes at greater distances from the base. Thus, the probability of detection increases the closer REDFOR agents get to the BLUFOR base (Figure 21), thereby making it more advantageous to strike as soon as possible.



Black dots represent REDFOR agents detected and removed from the game. The figure is representative of most games, indicating the low likelihood of REDFOR getting close to the BLUFOR base before they begin shooting and getting detected.

Figure 21. REDFOR Probability of Removal

[131] Fang Qiwan, Yin Zhixiang, and Jiang Chuanfu, *Menace of Anti-Ship Missiles and Shipborne Laser Weapons*, NAIC-ID(RS)T0337-96 (Wright-Patterson AFB, OH: National Air Intelligence Center, 1996), http://www.dtic.mil/dtic/tr/fulltext/u2/a313312.pdf.

Finally, the value for *maximum agent delay time* (0.006) indicated that launching the REDFOR attack as soon as possible generated the highest average base damage. By nearly minimizing this delay, REDFOR was able to start positioning its agents within firing range before BLUFOR was fully established. This makes sense considering that BLUFOR staggers the initial deployment of its drones. Figure 22 depicts this advantage by showing that when REDFOR begins its early attack, BLUFOR is still deploying its force, with only fifteen of its fifty agents on the board at the time.

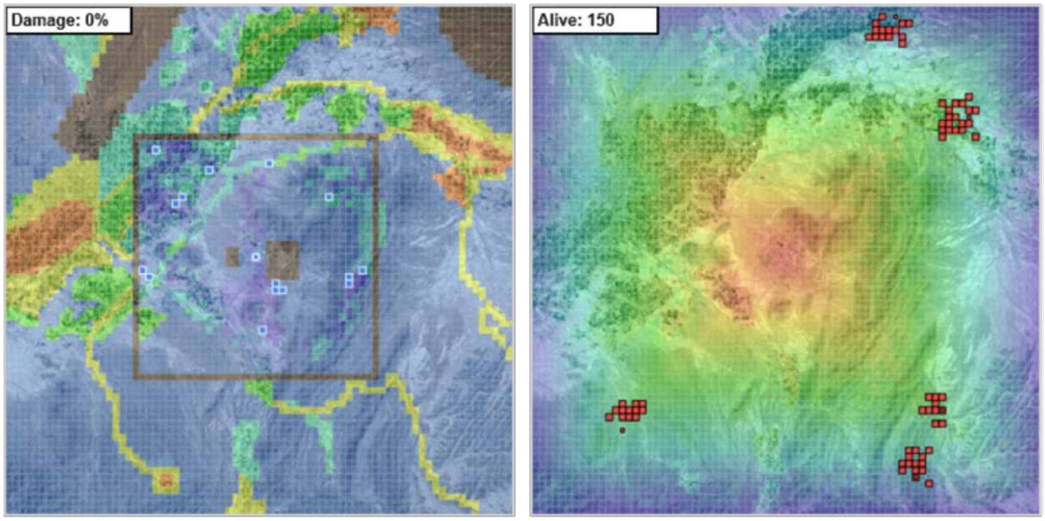


Figure 22. Example BLUFOR Disposition for Early REDFOR Attack

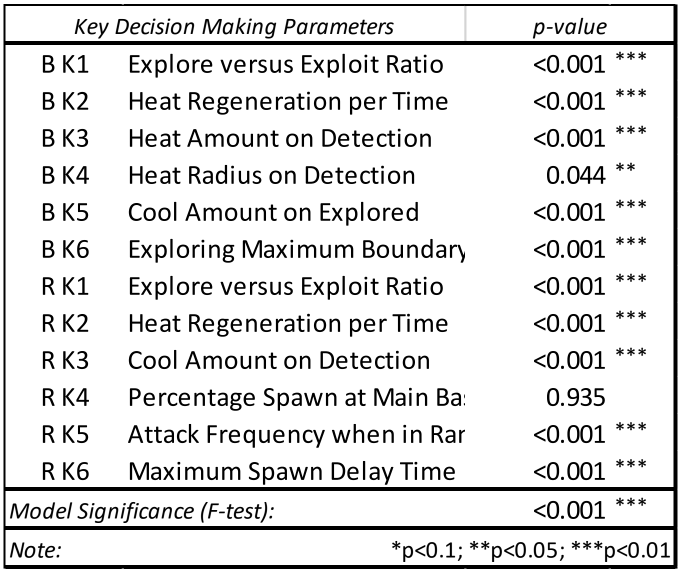
Ultimately, due to the degree of human bias in the development of the wargame, this research was able to analyze the resultant key decision-making parameters of the swarm logically. Explaining why the framework did what it did increases the level of trust that operators, within the human-autonomy team, will have when deploying a swarm. This insight frames the left and right limits of the swarm's behavior while accomplishing a specific mission objective. Additionally, this highlights a strength of the framework, in that it reduces the "black-box" effect commonly associated with AI systems. By focusing on a specific mission and with understandable decision-making parameters, this framework enables a process for establishing a higher degree of trust in the final solution.

3. Statistical Analysis

Statistical analysis of the underlying model and the resulting swarm algorithm effectiveness provides additional insight that can increase both trust and confidence in the solution that the training framework generated. This section examines three statistical tests that provide context into the validity and further utility of the experimental results. First, this section covers a regression analysis between the agent key decision-making parameters (factors) and the average base damage (dependent variable). Second, this section conducts a regression analysis to highlight residual risk inherent in the trained solution produced by the chosen assumptions. Third, this section presents a difference in means hypothesis test comparing the final algorithm effectiveness against the reference baseline algorithm.

The personality polygons after training are both distinct and evolutionarily stable, supporting the claim that all key decision-making parameters for BLUFOR and REDFOR had a significant impact on the amount of average base damage that occurred. To verify this initial observation, and to further understand which key decision-making parameters had the most influence, this analysis performed a statistical regression between the independent factors and dependent variables. Table 7 summarizes the regression results.

Table 7. Statistical Significance of Key Decision-making Parameters

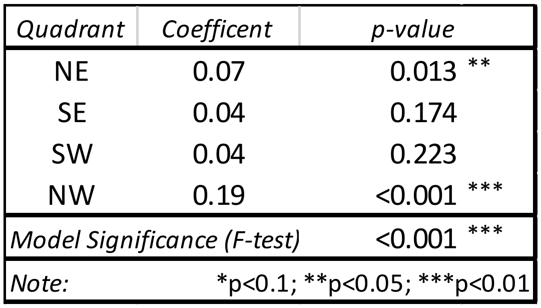


Key Decision Making Parameters		*p-value*	
B K1	Explore versus Exploit Ratio	<0.001	***
B K2	Heat Regeneration per Time	<0.001	***
B K3	Heat Amount on Detection	<0.001	***
B K4	Heat Radius on Detection	0.044	**
B K5	Cool Amount on Explored	<0.001	***
B K6	Exploring Maximum Boundary	<0.001	***
R K1	Explore versus Exploit Ratio	<0.001	***
R K2	Heat Regeneration per Time	<0.001	***
R K3	Cool Amount on Detection	<0.001	***
R K4	Percentage Spawn at Main Ba	0.935	
R K5	Attack Frequency when in Rar	<0.001	***
R K6	Maximum Spawn Delay Time	<0.001	***
Model Significance (F-test):		<0.001	***
Note:		*p<0.1; **p<0.05; ***p<0.01	

The results of this statistical analysis largely supports the claim that the chosen model parameters were significant factors, with almost every factor across BLUFOR and REDFOR showing a high level of statistical significance ($p<0.01$) in how they affect the percentage of base damage. However, there are two exceptions. First, although the *heat radius on detection* for BLUFOR displayed statistical significance ($p<0.05$), it was not as significant as the other factors for BLUFOR. This could be because another factor, like *heat amount on detection*, could interact with this term and reduce its impact. Next, the *percentage spawn at main base* parameter did not display statistical significance ($p>0.9$). Even though the trained value converged at 0.299, statistical regression did not indicate that this value had strong influence on the average base damage. This could be because the wargame did not give REDFOR the choice of how many locations it could spawn from, but rather how many forces it consolidated into its main effort. Also, the wargame did not allow different spawn locations to stager their attacks, relative to one another. These facts likely diminished the impact this factor had on the average base damage.

Even though the final BLUFOR personality polygon represents the best-trained swarm algorithm, the framework also offers the ability to identify residual risk inherent in the specific scenario. By examining the final matchups between the trained BLUFOR genomes and the MDCOA REDFOR genomes, analysts can determine common aspects within the sample data that lead to the worst outcomes. For instance, a particular portion of the environment may offer greater advantages to the attacking force, which then consistently leads to higher average base damage, regardless of the defending swarm's algorithm. To determine if this was the case, the analysis included a statistical regression between the number of spawn points within one of the four main quadrants of the board (northeast, southeast, southwest, and northwest) and the overall base damage. Table 8 depicts the results of this regression. The data for this regression came from a sample of twenty iterations of the wargame between the trained BLUFOR and the baseline REDFOR genomes.

Table 8. Regression Analysis to Highlight Residual Risk



Quadrant	*Coefficent*	*p-value*	
NE	0.07	0.013	**
SE	0.04	0.174	
SW	0.04	0.223	
NW	0.19	<0.001	***
Model Significance (F-test)		<0.001	***
Note:	*p<0.1; **p<0.05; ***p<0.01		

The results of this analysis indicated that REDFOR has a statistically significant advantage by spawning more forces from the northwest portion of the board (*p*<0.01), with a nearly three-fold increase, on average, in the amount of base damage produced when compared to the next highest region. This is likely because the terrain complexity in this portion of the board gives REDFOR an advantage by making it more difficult for BLUFOR to find them. Highlighting this residual risk allows commanders to design a layered defense, focusing efforts to the northwest, that combines other countermeasures with a swarm. This increases confidence that the swarm will produce the intended results.

Finally, this research performed a difference in means hypothesis test comparison between the performances of subsequent generations of the swarming algorithm against the reference baseline algorithm. First, the initial population of BLUFOR genomes created using the quasi-random process played the reference baseline REDFOR. The results of this run provided a starting point to compare subsequent generations of the BLUFOR algorithm, which resulted in an average base damage just over 120%. Next, the elite BLUFOR produced from the first generation played against the baseline REDFOR to produce the second point, resulting in an average base damage of 70%. The process accomplished baseline checks every five generations until the training cycle concluded. After thirty generations, a difference in means hypothesis test compared the resulting average base damage of 41% to the initial (pre-trained) baseline algorithm effectiveness. The test concluded at a 95% confidence level that, against the most-dangerous REDFOR, the difference between the average base damage from the initial generation (pre-training) and final generation (post-training) of BLUFOR is a reduction of 78-82% damage.

V. CONCLUSION

The experimental results from this thesis show that machine learning techniques are effective at training the decision-making algorithms for a decentralized swarm of drones using the rapid self-play of a mission-specific wargame. Although these results offer a promising indication of the value of the proposed Way of Swarm framework, the findings also raise two additional questions: (1) where can the military use the framework to enhance its forces (i.e., “so what”), and (2) what are additional research efforts that can improve the framework (i.e., “now what”)? The conclusion explores these questions by examining where this proof-of-concept model could scale up to meet the requirements of American military forces that would benefit by deploying decentralized swarms across all battlefield domains. Additionally, the conclusion discusses how further investments of resources and research into the framework components, such as improved model designs, sensitivity analysis, field tests, and user interfaces, can extend this research to serve as a stronger foundation for an operationalized framework for training America’s swarms.

The following chapter is presented in four sections. First, the chapter reflects on three major findings from this thesis to reinforce the key lessons learned from swarming research and experimentation with the custom-built swarm model. Second, the chapter reviews why selecting decentralized drones was an important factor for this research and it presents a set of mission-specific vignettes to highlight cross-domain applications for training decentralized swarms. Third, the chapter offers a series of further research opportunities that can improve the proposed framework, increase the trust for operationalizing the resulting algorithms, and enhance the overall understanding of when decentralized (versus centralized) control of swarming autonomous systems is more effective for a variety of mission types and environments. Finally, the chapter concludes with an inclusive summary of the thesis that reemphasizes the relevance and urgency of swarming research and reiterates how the role of the human decision maker is shifting—but is no less critical—in drone swarm ways of warfare.

A. MAJOR FINDINGS

1. Algorithms should be a Training Priority

Beyond merely focusing on training the human operator to be better at controlling their machines, the American Way of War must adapt to prioritize training the decision-making algorithms of the machines. Traditionally, military leaders design robust training programs with the intent to improve the decision-making process of their personnel and to enhance their effectiveness at accomplishing a specific task or mission.[132] However, with the emergence of AI and machine learning techniques, leaders must begin to adjust their long-held paradigms of what it means to train the operational forces; they must consider that the algorithms inside the machines can also improve their decision-making process by conducting similar robust training programs to those of personnel.

By developing training programs for machines, there is no longer a requirement for the algorithms for an autonomous system to be hard-coded by a human programmer (i.e., prescribing exactly what parameters the machine should use to make its decisions). Instead, it is possible to develop a shell of an algorithm, without knowing optimal decision-making parameters, and to allow machine learning to develop a recommended set of parameters that is best fit to achieve a specified task. For this thesis, the proof-of-concept model demonstrated that this type of machine learning framework is possible for training decentralized swarms, and with additional research, this process may prove to be faster and more successful at optimizing solutions compared to a human programmer.[133] Therefore, with a proper algorithm training framework, the idea of drones "self-learning" how to accomplish a mission as a collective could become just as routine as personnel who learn from their hands-on experiences (similar to reinforcement machine learning) or learn from a subject matter expert (similar to supervised machine learning).

[132] J. Fletcher and P. Chatelier, *An Overview of Military Training*, IDA Document D-2514 (Alexandria, VA: Institute for Defense Analyses, 2000), http://www.dtic.mil/dtic/tr/fulltext/u2/a408439.pdf.

[133] Silver et al., "Mastering the Game of Go Without Human Knowledge," 2017; Vincent, "AI Bots Beat Humans," 2018.

Optimizing the decentralized decision-making algorithms of autonomous systems is essential for overcoming the span-of-control limitations for centrally coordinating large quantities of drones. The ultimate objective of the human-autonomy team is to optimize the combined decision-making effort of the human *and* the machine. Furthermore, when the human decision maker becomes over-tasked or is out-performed in their mental or physical reaction times, then the team must prioritize building trust and confidence in the algorithms of the autonomous systems in order to enable delegating more of the cognitive workload. Once the machine is proficient at tasks, below the level of accepted risk where the operator trusts the actions of the algorithm, then the operator can assign more tasks to the machine. This rebalances the workload of the team, allowing the operator to focus on mission aspects that are unique to the person (e.g., ethics, empathy, risk). Ultimately, it is not the decision-making process (i.e., OODA loop) of either the human *or* machine, that must outperform the adversary's decision-making process, but it is the combination of the overall human-autonomy team that military leaders must effectively train to succeed.

2. Design Wargames for Machine Learning Integration

The right type of wargame is necessary to serve as a training platform to generate a large quantity of datasets for machine learning techniques to train swarming algorithms. Currently, there are no prerecorded datasets that demonstrate how a drone swarm should optimally behave in different military mission sets or environments. Thus, a framework that uses simulated agents and environments to generate data is suited to the data-sparse field of swarm research. However, to generate millions of game results for reinforcement machine learning techniques to learn from those experiences, the wargame design needs to be capable of fully executing simulated missions in a short time span. There are very few existing wargames that are designed to rapidly execute millions of games and that are built to integrate machine learning with direct access into the parameters of wargame agents. To overcome this limitation, and to generate higher quantities of game results, designers should ensure that the wargames are capable of executing at a speed faster than real time, that wargames can be entirely self-played by AI players, and that wargames can distribute across multiple computers that collaborate in data collection and generation.

Furthermore, wargame designers for swarm algorithm design should consider the benefits of resisting the trend of designing software with immersive real-time interfaces (e.g., touchscreen or virtual reality) and resisting a priority on high-definition graphics.[134] While these types of simulations can be ideal for enhanced human training, the delays for constantly waiting for mandatory user inputs and the processing requirements to display real-time graphics reduces the pace for an AI system to rapidly iterate through millions of self-played games. Therefore, as designers consider how to incorporate the power of machine learning techniques to determine drone agent decision-making parameters, they should consider how their wargames can benefit swarm algorithm design through either faster game play (for the machine) or higher-fidelity game play (for the human).

3. Machine Learning is Effective and Accessible

One of the most significant findings from this thesis research was that machine learning techniques and the computing power necessary to effectively train the decision-making algorithm for a swarm of autonomous systems are readily accessible resources. The genetic algorithm that was effective in this research at rapidly searching through a vast solution space for optimized parameters is reproducible in any coding language. Additionally, companies such as Google, Facebook, and IBM are packaging and freely distributing their more powerful machine learning techniques (e.g., deep artificial neural networks and Q-learning algorithms).[135] Furthermore, with the decreasing costs of data storage, and the increasing speeds of networking, there exists low-cost solutions to create highly capable data centers and distributed cloud-computing environment that can extend beyond large companies or research institutions.[136] This emerging combination of readily available access to data processing power, data storage, and data-driven machine learning

[134] Philip Sabin, *Simulating War: Studying Conflict through Simulation Games* (London, UK: Continuum International Publishing Group, 2014), 35.

[135] Murray Newlands, "The Democratization of Machine Learning is at Hand and this AI Company is on the Front Lines," *Forbes,* October 20, 2017, https://www.forbes.com/sites/mnewlands/2017/10/20/the-democratization-of-machine-learning-is-at-hand-and-this-ai-company-is-on-the-front-lines/#32586ede2451.

[136] Nabil Sultan, "Cloud Computing: A Democratizing Force?" *International Journal of Information Management* 33, no. 5 (October 2013), https://www.sciencedirect.com/science/article/pii/S0268401213000820.

techniques is revolutionizing the ability to transition from studying the theories of what is possible through AI advancements, and applying those theories to real challenges.

An important concept explored in this thesis research is that the computing power of large cloud-computing data centers does not need to be inside of the relatively small hardware of each drone agent to harness the benefits of machine learning techniques. Moreover, there does not need to be a stream of communication between the swarm and the cloud, since each drone does not need to wait on a centralized controller or algorithm to decide what action each agent should make. Instead, this framework demonstrated that operators can use cloud-computing services, narrow AI, and wargame simulations *before* the mission to train a predetermined algorithm set for each agent. The operator can then copy and upload the trained parameters (a smaller computer file that is computationally inexpensive to run) into the individual drones to execute. This process of pretraining the algorithm and then deploying it is similar to how cellular phones can learn to recognize faces or voices in their deployed applications.[137] Although the algorithm of how to detect specific features was pretrained on powerful data center computers, once the algorithm was trained, it can deploy to less powerful computers for execution.

B. OPERATIONALIZING THE FRAMEWORK

The emergence of commercially available technologies, like AI and autonomous systems, is progressing warfare into its next evolution, where any force, including both non-state and state actors, can deploy a swarm into combat. Given this change in the character of war, figuring out how to not only leverage swarms, but also how to defend against them, is critical for a competitive strategy. Many within the DoD have recognized the rising cost of military hardware and the corresponding decrease in the quantity the services can field.[138] The military has sought to reverse this trend by investing in large quantities of cheap systems (like drone swarms), yet it has overlooked a crucial element:

[137] "Machine Learning and Mobile: Deploying Models on the Edge," *Algorithmia*, June 21, 2018, https://blog.algorithmia.com/machine-learning-and-mobile-deploying-models-on-the-edge/.

[138] Martin Edmonds, "Augustine's Laws: Norman Augustine," *Defense and Security Analysis* 21, no. 1 (2005): 111-114, https://www.tandfonline.com/doi/pdf/10.1080/1475179052000341542.

how to train this autonomous swarming component to perform the variety of different missions across the services. To leverage the full potential of swarms, the DoD must focus algorithm training on specific missions and train to decentralize swarm execution.

Although centralizing the control of swarms works well in certain circumstances, it may not necessarily work in all operating environments. Decentralized swarms are more resilient, given that centralized execution presents a single point of vulnerability that could be problematic in the contested environments of near-peer adversaries. In addition, decentralized swarms provide versatility by enabling faster responses to a variety of dynamic circumstance, thereby minimizing the span-of-control challenges. Developers of naval doctrine have highlighted the value of being able to operate along this spectrum of command and control to full autonomy, or a spectrum between network-centric warfare and network-optional warfare.[139] Decentralization of swarms presents a difficult challenge, but one that the American Way of War must endogenize to enhance mission performance in dynamic environments and to overcome adversary actions that seek to disrupt, deny, and degrade America's current advantages.

To operationalize the framework presented in this thesis, the DoD must undertake an effort to create a variety of different wargames that machine learning techniques can self-play and master. To obtain the benefits of resiliency and versatility of a decentralized swarm, trust must exist within the human-autonomy team. Ultimately, building trust in the tactical environment comes down to higher-quality training. Therefore, establishing enough trust to decentralize decision making from the human to the individual drones within a swarm requires a novel approach to training. There are numerous different mission sets across all domains that could benefit significantly from this thesis initiative. The following section explores two relevant vignettes that highlight this potential.

[139] Christopher Nelson, "Fleet Tactics Returns – A Conversation with Authors Wayne Hughes and Bob Girrier," *Center for International Maritime Security*, July 30, 2018, http://cimsec.org/fleet-tactics-returns-a-conversation-with-authors-wayne-hughes-and-bob-girrier/37040.

1. Vignette 1: Unmanned Undersea Vehicle Patrols

The Navy is seeking to leverage emerging undersea systems, like Hydroid's Remus M3V and Teledyne Energy's subsea drone refueling stations, for search, survey, and reconnaissance operations.[140] In one scenario, the Navy envisions deploying swarms of unmanned undersea vehicles (UUVs) ahead of a fleet operating in contested waters to survey the area for threats. Research sponsored by the Navy that examines how to employ UUVs for these missions claim that *assured communications* is critical to control their systems.[141] This is because unmanned systems still need human operators and will thus "exacerbate manpower and manning challenges... [and] in many instances, the number of personnel required to operate and support a single unmanned system exceeds that for a manned platform with a similar concept of employment."[142] This approach, which relies heavily on the communications network, would centralize the operation of these drones under the Navy's existing command and control architecture. In a world of perfect communication, a network-centric strategy like this may produce a high probability of mission success.

However, the enemy gets a vote, and in contested environments, like the South China Sea, the Navy cannot assume a threshold level of communication within its entire network. To improve the ability to operate UUVs in challenging environments, the Navy needs to embrace a *network-optional approach* and decentralize its underwater swarms to increase both resiliency and versatility.[143] To meet this need, the Navy should invest in the development of a mission-specific wargame that a narrow AI could self-play to train the local decision-making parameters of the individual UUVs and refueling stations. Much like the model presented in this thesis, the result of this effort would be a highly

[140] Victoria Leoni, "New Undersea Drones are Smaller, Cheaper, and Can be Refueled Deep Under Water," *Defense News*, April 10, 2018, https://www.defensenews.com/digital-show-dailies/navy-league/2018/04/10/new-undersea-drones-are-smaller-cheaper-and-can-be-refueled-deep-under-water/.

[141] Scott Savitz et al., *U.S. Navy Employment Options for Unmanned Surface Vehicles*, RR-384-NAVY (Santa Monica, CA: RAND, 2013), xxvii, https://www.rand.org/content/dam/rand/pubs/research_reports/RR300/RR384/RAND_RR384.pdf.

[142] Ibid, xxix.

[143] Nelson, "Fleet Tactics Returns," 2018.

specific (but adaptable) algorithm that the Navy could use to decentralize its swarms of UUVs to accomplish the search, survey, and reconnaissance mission. Importantly, this does not imply that there should not be an *option* to centralize or override the operation of their UUVs. Yet, if the Navy cannot transmit commands due to a need to limit electronic emissions, if the swarm cannot receive commands due to adversarial jamming, or if the scale of the swarm outpaces available manpower, then decentralization provides a way to overcome these obstacles and ensure freedom of action.

2. Vignette 2: Collateral Scans for Kinetic Strikes

Air strikes against terrorist leaders and other high profile figures within violent extremist organizations have become a mainstay of the American strategy to defeat their spread.[144] Oftentimes, the DoD uses the term *kinetic strike* to describe these operations, which refers to "lethal air action controlled by dislocated strike cells against enemy time-sensitive targets, and/or high-value individuals (HVI)."[145] Although military analysts argue that kinetic diplomacy is never sufficient, they concede that an effective strategy still requires some level of violence, albeit discriminate violence.[146] Therefore, increasing the performance of kinetic strikes is beneficial to the National Defense Strategy.

However, executing these tactics is resource intensive. Often the demand for assets with the requisite combination of training, sensors, endurance, and munitions, outpaces the available supply. One particular resource-intensive task within the kinetic strike mission is conducting real-time *collateral scans*, or scanning the area to reduce the risk to non-combatants. Typically, multiple assets with visual sensors perform this task by scanning the target area of interest for any collateral concerns (e.g., vehicle traffic,

[144] Dan De Luce and Sean Naylor, "The Drones are Back," *Foreign Policy*, March 26, 2018, https://foreignpolicy.com/2018/03/26/the-drones-are-back/; Daniel Rosenthal and Loren Schulman, "Trump's Secret War on Terror," *Atlantic*, August 10, 2018, https://www.theatlantic.com/international/archive/2018/08/trump-war-terror-drones/567218/.

[145] Michael Smith, "Kinetic Strike for Special Tactics to Achieve Precision Strike Effects" (presentation, U.S. Air Force Weapons School, Nellis AFB, NV, 2016), https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2016/psts/CaptMichaelDSmith.pdf.

[146] Monica Toft, "The Dangerous Rise of Kinetic Diplomacy," *War on the Rocks*, May 14, 2018, https://warontherocks.com/2018/05/the-dangerous-rise-of-kinetic-diplomacy/.

pedestrians, etc.). Determining the locations in which the aircraft should focus its scans and how to timely and accurately communicate potential collateral concerns in a rapidly-changing environment are difficult tasks that are predominantly centralized.

To increase the performance of kinetic strike missions, the DoD should develop an on-demand swarm of decentralized drones to perform the collateral scan mission. For example, using the Perdix micro-drone test[147] as a baseline, existing aircraft would deploy a swarm of small, cheap, and lightweight drones to conduct collateral scans when conditions were met for a kinetic strike. This would decrease the number of traditional assets needed to mitigate the risk of collateral damage, and replace them with a resilient and versatile swarm of drones. Operating in this manner would require trust within the human-autonomy team. To achieve this, the DoD needs to develop a collateral scan wargame that a machine learning algorithm could self-play to train the local decision-making parameters of the individual drones. In operation, the appropriate authority would first identify and pass the intended target to the trained swarm before deployment. Once deployed, the swarm would interact with the local environment, based on its AI assisted and validated training, to determine the best locations to position themselves, identify collateral concerns, and communicate concerns in a timely manner.

Decentralizing the operation of drones performing collateral scans during a kinetic strike does not imply that the human is removed from decisions to employ lethal force. Rather, the aim is to speed the decision-making process of the human-autonomy team by allowing the swarm to self-organize and position itself at the best locations to conduct collateral scans and streamline real-time adjustments in a dynamic environment. Trying to centralizing the operation of a swarm in this type of scenario would at best be inefficient, and, at worst, be counterproductive and could result in missed opportunities to conduct a successful kinetic strike. Ultimately, applying the proposed framework of this thesis to the kinetic strike mission has the potential to produce swarm strategic utility.

[147] Department of Defense, “Perdix Fact Sheet,” 2018.

C. FUTURE RESEARCH

There are multiple approaches to advance the research and experimentation of the proposed Way of Swarm framework. This section highlights several areas of suggested future research, including improvements to the model, applications of sensitivity tests for risk mitigation and hardware investment priorities, opportunities for field-testing trained algorithms, innovations for building user interfaces, and experiments to test the spectrum of control for autonomy based on mission sets and performance (i.e., what conditions are best to centralize as opposed to decentralize the control of a swarm).

1. Improve the Model

Future researchers can improve upon the proof-of-concept model in this thesis by addressing both the wargame design and the machine learning techniques selected. For instance, researchers could enhance the wargame to simulate the specific capabilities of a known drone type, model a specific environment with greater accuracy, or reduce the level of abstraction in the wargame. For instance, adding more fidelity to the wargame, such as three-dimensional flight characteristics, adverse weather effects, and line-of-sight communications, will allow the machine learning technique to solve for a greater variety of unknown key decision-making parameters (e.g., optimal altitudes for flying drones). However, the more robust the simulation, the more time it will take to execute millions of simulations to generate datasets. Researchers will need to determine the right balance between the fidelity of the wargame, the desired key decision-making parameters, and the amount of resources available to iterate through millions of simulations.

Additionally, researchers can incorporate more sophisticated machine learning techniques, such as artificial neural networks or Q-learning, to solve for decision-making parameters at a broader range. The underlying decision-making algorithm for the drone agents in this thesis was based on a set of linear equations containing a series of rates used for heating and cooling gameboard tiles along a critical decision-tree path. The genetic algorithm systematically adjusted the rates and proportions to determine optimal global behaviors. Alternatively, neural networks offer a design structure for discovering this decision-making algorithm by reducing the human template (i.e., the OODA loop) down to a more primitive set of raw inputs and outputs. This approach could potentially reduce the amount of human investment when scaling this model across multiple military mission sets (e.g., UUV patrols, collateral scan for kinetic strikes, etc.), but it requires more research to optimize neural network designs for use in decentralized applications.

Moreover, *Q-learning* is another powerful version of reinforcement learning that could be used to enhance the framework design. Q-learning uses a values function within the algorithm to choose an action based on a given state to earn an immediate and future reward. When the possible combinations of actions and rewards is large, a Q-learning can be used to output and update value functions as opposed to individual agent actions.[148] Therefore, as the complexity of the wargame increases, Q-learning provides a method to "determine the best weightings for optimal control and design problems."[149] Overall, employing additional machine learning techniques could make the final swarm training framework more efficient and represents a valid area for further research.

2. Conduct a Robust Sensitivity Test

Creating both trust and confidence within the human-autonomy team is a critical objective to bring more autonomy to the battlefield. Chapter IV, Section B, discussed one method for using the framework to highlight residual risk via statistical analysis of the final trained algorithms. However, in addition to these tests, the framework needs a more robust series of tests to further understand the sensitivity of the wargame model and initial assumptions. Once the training framework generates a solution for a given mission and scenario, additional iterations should examine the implications of any errors in the assumptions used to develop the wargame. For instance, if the designers of the wargame underestimated the adversary’s capability (e.g., speed, weapons range, etc.) then the trained decision-making parameters of the swarm would not perform as advertised.

[148] Volodymyr Mnih, et al., “Human-Level Control through Deep Reinforcement Learning,” 2015.

[149], Kaivan Kamali et al., “Using Q-Learning and Genetic Algorithms to Improve the Efficiency of Weight Adjustments for Optimal Control and Design Problems,” *Journal of Computing and Information Science in Engineering* 7, no. 4 (December 1, 2007): 302–308, http://computingengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1400904.

Hence, additional sensitivity tests provide a way to better understand the risks inherent in the solution, which increases both trust and confidence for the operator.

Running a robust sensitivity test would consist of three distinct steps. First, designers of the wargame would modify key simulation assumptions. For example, if the original training iteration assumed an enemy attack range of one kilometer, designers would modify this value (e.g., increase to two kilometers). Second, developers would re-run the training framework with the adjusted assumption to observe the differences in swarm algorithms produced. By changing the enemy's characteristics, the drone swarm will respond by adapting its algorithm over the course of the training evolution, resulting in a different solution. Third, to understand the sensitivity of the model to each particular assumption, the test would need to culminate with a series of wargames between the new optimized adversary produced with the adjusted assumptions and the original drone swarm algorithm. This final step provides insight into the changes in the results that would highlight the risk inherent in each assumption used to build the model.

Model sensitivity tests provide an analysis tool for commanders to further identify residual risk and to gain trust and confidence in the framework. Commanders deploying swarms that trained through a narrow AI can use this tool to determine ways to improve the assumptions within their control (e.g., shape their critical information requirements). This can improve the training model, and thereby improve outcomes on the battlefield. Not only can this tool identify which characteristics of the adversary present the greatest risk, but the tool can also indicate which characteristics of the drone swarm are the most critical for additional resource investment (e.g., sensor upgrades, battery life, speed, etc.). Furthermore, commanders could use this tool to determine how many drones it would take to achieve a desired mission effectiveness percentage or an acceptable level of risk.

3. Field-Test the Trained Algorithms

Although the experimental results support that the proposed training framework produces effective drone swarm algorithms, without any real-world flight validation, the operators responsible for mission accomplishment would likely be averse to trusting the swarm and employing them in combat. Field tests of the machine learning recommended algorithms are the final step to build trust in the decision-making process of the swarms and serve as a location to build partnerships in the resulting human-autonomy teams. Installations, such as Camp Roberts in California, currently exist to field-test aerial drone swarm tactics, and organizations like the Naval Postgraduate School, Georgia Tech, and DARPA have used the installation extensively for the purpose of live flight validation.[150]

Field tests can generate data to reduce the gap between the theoretical simulation and reality. Comparing real-world results to the wargame simulation provides the human and narrow AI with feedback to update the assumptions in the model. The model relies on updated assumptions to generate refined behaviors for the swarms in both subsequent training and operations. Therefore, combining additional sensitivity tests in simulation, along with real-world operational testing is critical for culminating a viable framework that remains rapid, agile, and flexible.

4. Design a User Interface

One of the critical areas of further research is determining the best practices for constructing a set of user interfaces for human operators to monitor and override drone swarms (a front-end interface) as well to change assumptions and mission-types for the algorithm model and to retrain key decision-making parameters (a back-end interface). For the front-end interface, the software package must be intuitive to deploy and operate. If the swarm has real-time communication links available during operations, the interface should also include options to monitor active swarm operations and to update information to the swarms with critical changes (e.g., updated intelligence priorities). Additionally, the front-end interface should have the option to manually override specific drone agents when deemed necessary by a human operator. Although the intent of the framework is to train the swarm to effectively operate without the need for constant human control, due to unforeseen circumstances, the consideration must be assessed to keep an override option available for the operators to switch to manual control mode as desired.

[150] "Service Academies Swarm Challenge Live-Fly Competition Begins," DARPA Outreach, April 23, 2017, https://www.darpa.mil/news-events/2017-04-23; U.S. Air Force, Autonomous Horizons, 2015.

Beyond just being able to launch, monitor, and override drone swarms in the field, one goal of the swarm training framework is that the tools for modifying the wargames, updating mission assumptions, and retraining the swarms should also be accessible at forward locations. Being able to rapidly adjust the training of drone swarms to current scenarios and assumptions that are known by the operators of the swarms in the field will give those swarms a higher effectiveness than those swarms trained for more generalized missions and environments. However, this goal requires that researchers build and test a back-end user interface that allows users who are not computer programmers or machine learning experts to apply the framework to their unique drone swarm and mission set. The back-end interface should allow users to easily change the type of drone agents, the environmental parameters, intelligence assumptions, and the mission-specific objective of their assigned swarm, and then click "train" to output an updated and optimized set of decision-making parameters that the framework generates to fit a specific situation.

5. Wargame the Spectrum of Control

One critical claim made in this thesis research is about the importance of training decentralized (as opposed to centralized) swarms of autonomous systems. Future research could further explore which control modes are more effective for military drone swarms in unique mission sets and environmental conditions. In addition to testing whether one mode of control is better in certain situations, experimentation could also help determine best practices for building a hybrid framework of control. Is it always best to employ swarms under centralized control as long as the swarm can maintain communication with a powerful data center? Or, perhaps, is it best for swarms to execute as much as possible in a decentralized pre-trained algorithm mode, even if they have good communications, so that military personnel will have more opportunities to build trust on how the swarm responds in a variety of situations? In either case, as military forces train to fight in contested and degraded environments, the ability to survive without communication, whether intentional (i.e., network-optional) or unintentional (i.e., the enemy gets a vote) demands having the best-trained decentralized swarming techniques ready for forces to execute when necessary.

To emphasize the importance between decentralized and centralized control for drone swarms, future wargame designers should add the ability for narrow AI to assume different modes of control over individual drones or entire forces in their wargames. There is a lack of wargames that focus on how to better improve the decentralized control of agents, as opposed to focusing on improving the single human (or computer) player who usually directly controls all agents in the game. Instead of insisting on centralizing the game information to a single controller, more wargames should focus on how the agents could improve their decision-making ability given the agent's constrained observations and understanding of the state of the game. A wargame can simulate the demand for decentralization by limiting agents' "knowledge" of the state of the game so that not all agents always share perfect information. Essentially, each agent in a wargame gets its own observations and decision-making algorithm (i.e., its own OODA loop) and cannot always reliably receive new commands. With more focus on the agents, operators will concentrate less on maximizing their own decision-making processes, and instead, become more cognizant of the benefits of optimizing the agents' algorithms.

D. SUMMARY

Ultimately, as the ability to employ small, inexpensive, and capable autonomous systems is increasing across all domains, the American military needs to adapt its way of fighting to ensure it maintains a competitive advantage. The current frameworks that intend to train swarm tactics to drones limit the potential for the human-autonomy team to succeed in complex, contested, or denied environments, by relying on communication to a centralized controller that will begin to over-task the human decision maker. To overcome this span-of-control barrier, the proposed Way of Swarm is a successfully demonstrated method to develop decentralized swarming algorithms based on training the decision-making parameters of individual drone agents. The synergy of mission-specific wargames, machine learning, and cloud-computing services provided a rapid, agile, and flexible framework to train swarming agents and generate effective training datasets in an otherwise data sparse area of research. When combined with future research that includes better models, a robust sensitivity analysis, field tests, and accessible user interfaces, the proposed framework will serve as a force multiplier to enhance human-autonomy teams.

The proposed Way of Swarm aims to enhance the human-autonomy team, but not completely remove the human from the decision-making process. Although this research trained a decentralized swarm to better accomplish a mission with minimal human input, this does not mean that human input and oversight is not a requirement. Rather, the goal of a decentralized drone swarm is to enable individual agents to make their own decisions given the operator's intent, as part of a collective team, thereby rebalancing the workload of the human. This enables the operator to focus on tasks that are not suitable for AI-trained systems, while maximizing the unique capabilities of the human in the team. This falls in line with the DoD's vision for unmanned systems, contending that "the expansion of capabilities in unmanned systems over the coming decades will largely be dependent on the ability to effectively team humans and autonomous systems in the force."[151] This thesis has argued that effective teaming between human operators and swarms of autonomous systems requires trust in decentralized execution; building this trust is achieved by adapting new technologies into the strategic framework for training effective algorithms.

[151] Department of Defense, *Unmanned Systems Integrated Roadmap FY 2017-2024*, (Washington, DC: DoD, 2017), 21, http://cdn.defensedaily.com/wp-content/uploads/post_attachment/206477.pdf.

LIST OF REFERENCES

Albers, Paul, and Han Vries. “Elo-rating as a Tool in the Sequential Estimation of Dominance Strengths.” *Animal Behavior* 61 (2001): 489-495. http://dx.doi.org/10.1006/anbe.2000.1571.

Algorithmia. “Machine Learning and Mobile: Deploying Models on the Edge.” June 21, 2018. https://blog.algorithmia.com/machine-learning-and-mobile-deploying-models-on-the-edge/.

Almond, Gabriel, and Stephen Genco. “Clouds, Clocks, and the Study of Politics.” *World Politics* 29, no. 4, 1977. https://www.jstor.org/stable/2010037?seq=1.

Arquilla, John, and David Ronfeldt. *Swarming and the Future of Conflict*. DB 311-OSD. Santa Monica, CA: RAND, 2000. https://www.rand.org/pubs/documented_briefings/DB311.html.

Badham, John, dir. *WarGames*. Los Angeles, CA: United Artists, 1983. http://www.netflix.com.

Bardenet, Rémi, Mátyás Brendel, Balázs Kégl, and Michele Sebag. “Collaborative Hyperparameter Tuning.” In *International Conference on Machine Learning* (2013): 199–207. http://proceedings.mlr.press/v28/bardenet13.pdf.

Barrett, Brian. “Inside the Olympics Opening Ceremony World-Record Breaking Drone Show.” *Wired*, February 9, 2018. https://www.wired.com/story/olympics-opening-ceremony-drone-show/.

Bianco, Tom. “GPU Acceleration Advancing the Evolution of Fast and Big Data.” *Datanami*, July 24, 2017. https://www.datanami.com/2017/07/24/gpu-acceleration-advancing-evolution-fast-big-data/.

Borshchev, Andrei, and Alexei Filippov. “From System Dynamics and Discrete Event to Practical Agent Based Modeling.” In *International Conference of the System Dynamics Society* (July 2004), 25–29. https://www.systemdynamics.org/assets/conferences/2004/SDS_2004/PAPERS/381BORSH.pdf.

Bousquet, Antoine. *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity*. London, UK: Hurst and Company, 2009.

Brooks, Rosa. *How Everything Became War and the Military Became Everything: Tales from the Pentagon*. New York, NY: Simon and Schuster, 2017.

Brown, Joel, and Thomas Vincent. “Organization of Predator-Prey Communities as an Evolutionary Game.” *Evolution* 46, no. 5, 1992. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1992.tb01123.x.

C., T. “What is the Singularity?” *Economist*, May 14, 2018. https://www.economist.com/the-economist-explains/2018/05/14/what-is-the-singularity.

Carbonell, Jaime, Ryszard Michalski, and Tom Mitchell. “An Overview of Machine Learning.” In *Machine learning: An Artificial Intelligence Approach*, edited by Ryszard Michalski. Berlin, Germany: Springer, 2013.

Cassano, Jay. “Pentagon’s Artificial Intelligence Programs Get Huge Boost in Defense Budget.” *Fast Company,* August 15, 2018. https://www.fastcompany.com/90219751/pentagons-artificial-intelligence-programs-get-huge-boost-in-defense-budget.

Chatham, Ralph E. “The 20th Century Revolution in Military Training.” In *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*, edited by Ericsson KA, 27–60. United Kingdom: Cambridge University Press, 2009.

Clark, Colin. “Defense Strategy Raises China to Top Threat; Allies Feature Prominently.” *Breaking Defense,* January 18, 2018. https://breakingdefense.com/2018/01/mattis-military-strategy-raises-china-to-top-threat-allies-feature-prominently/.

Coram, Robert. *Boyd: The Fighter Pilot Who Changed the Art of War.* New York, NY: Little, Brown and Company, 2002.

Cordesman, Anthony H., and Abraham R. Wagner. “The Lessons of the 1973 Arab-Israeli Conflict: October War.” In *The Lessons of Modern War: Volume 1: The Arab-Israeli Conflict, 1973–1989*, edited by Abraham R. Wagner, Boulder, CO: Westview Press, 1990.

Costlow, Terry. “How Big Data is Paying off for DoD.” *Defense Systems*, October 24, 2014. https://defensesystems.com/articles/2014/10/24/feature-big-data-for-defense.aspx.

DARPA. “OFFensive Swarm-Enabled Tactics (OFFSET).” Accessed September 20, 2018. https://www.darpa.mil/work-with-us/offensive-swarm-enabled-tactics.

DARPA. “Service Academies Swarm Challenge Live-Fly Competition Begins.” April 23, 2017. https://www.darpa.mil/news-events/2017-04-23.

De Luce, Dan, and Sean Naylor. “The Drones are Back.” *Foreign Policy*, March 26, 2018. https://foreignpolicy.com/2018/03/26/the-drones-are-back/.

Defense Information Systems Agency. “Cloud Services Support.” Accessed October 10, 2018. https://www.disa.mil/Computing/Cloud-Services/Cloud-Support.

Deneubourg, J. L., S. Aron, S. Goss, and J. Pasteels. “The Self-organizing Exploratory Pattern of the Argentine Ant.” *Journal of Insect Behavior* 3, no. 2, 1990. https://link.springer.com/article/10.1007/BF01417909.

Department of Defense. “Perdix Fact Sheet.” Accessed on February 11, 2018. https://dod.defense.gov/Portals/1/Documents/pubs/Perdix%20Fact%20Sheet.pdf?ver=2017-01-09-101520-643.

Department of Defense. *Autonomy in Weapon Systems*. DoD Directive 3000.09. Washington, DC: Department of Defense, 2017. http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/300009p.pdf.

Department of Defense. *Task Force Report: The Role of Autonomy in DoD Systems*. Washington, DC: Department of Defense, 2012. https://fas.org/irp/agency/dod/dsb/autonomy.pdf.

Department of Defense. *Unmanned Systems Integrated Roadmap FY 2017-2024*. Washington, DC: Department of Defense, 2017. http://cdn.defensedaily.com/wp-content/uploads/post_attachment/206477.pdf.

Department of the Air Force. *Risk Management (RM) Guidelines and Tools*. Air Force Pamphlet 90–803. Washington, DC: Department of the Air Force, 2013. http://static.e-publishing.af.mil/production/1/af_se/publication/afpam90-803/afpam90-803.pdf.

Department of the Army. *Commander and Staff Organization*. FM 6-0. Washington, DC: Department of the Army, 2017. http://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN7501_FM_6-0_Incl_C2_FINAL_WEB.pdf.

Department of the Army. *Military Decision-Making Process*. FM 101–5. Washington, DC: Department of the Army, 1997. http://www.au.af.mil/au/awc/awcgate/army/fm101-5_mdmp.pdf.

Department of the Army. *Operations*. FM 3–0. Washington, DC: Department of the Army, 2017. https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN6687_FM%203-0%20C1%20Inc%20FINAL%20WEB.pdf.

Diaz-Gomez, Pedro, and Dean Hougen. “Initial Population for Genetic Algorithms: A Metric Approach.” In *GEM* (2017): 43–49. http://www.cameron.edu/~pdiaz-go/GAsPopMetric.pdf.

Drew, Christopher. “Lockheed Lowers Price on F-35 Fighters, After Prodding by Trump.” *New York Times*, February 3, 2017. https://www.nytimes.com/2017/02/03/business/lockheed-lowers-price-on-f-35-fighters-after-prodding-by-trump.html.

Economist. “Getting to Grips with Military Robotics: Autonomous Robots and Swarms Will Change the Nature of Warfare.” January 25, 2018. https://www.economist.com/special-report/2018/01/25/getting-to-grips-with-military-robotics.

Edmonds, Martin. “Augustine's Laws: Norman Augustine.” *Defense and Security Analysis* 21, no. 1 (2005): 111-114. https://www.tandfonline.com/doi/pdf/10.1080/1475179052000341542.

Fehervari, Istvan, and Wilfried Elmenreich. “Evolving Neural Network Controllers for a Team of Self-Organizing Robots.” *Journal of Robotics*, vol. 2010, March 25, 2010. https://www.hindawi.com/journals/jr/2010/841286/abs/.

Fletcher, J., and P. Chatelier. *An Overview of Military Training.* IDA Document D-2514. Alexandria, VA: Institute for Defense Analyses, 2000. http://www.dtic.mil/dtic/tr/fulltext/u2/a408439.pdf.

Galdorisi, George. “Keeping Humans in the Loop.” *Proceedings* 141, no. 14, 2015. https://www.usni.org/magazines/proceedings/2015-02/keeping-humans-loop.

Garnier, Simon, Jacques Gautrais, Guy Theraulaz. “The Biological Principles of Swarm Intelligence.” *Swarm Intelligence* 1, no. 1, 2007. https://link.springer.com/article/10.1007/s11721-007-0004-y.

George Mason University’s Evolutionary Computation Laboratory. “MASON.” Accessed November 12, 2017. https://cs.gmu.edu/~eclab/projects/mason/.

Gianluca, Manzo. “Potentials and Limits of Simulation Multi-Agents: An Introduction.” *Revue Francaise de Sociologie* 55, no. 4, 2014, https://www.cairn-int.info/article-E_RFS_554_0653--the-potential-and-limitations-of.htm.

Gibbons-Neff, Thomas. “ISIS Drones are Attacking U.S. Troops and Disrupting Airstrikes in Raqqa.” *Washington Post*, June 14, 2017. https://www.washingtonpost.com/news/checkpoint/wp/2017/06/14/isis-drones-are-attacking-u-s-troops-and-disrupting-airstrikes-in-raqqa-officials-say.

Giles, Kathleen, and Kristin Giammarco. “Mission-based Architecture for Swarm Composability (MASC).” *Procedia Computer Science* 114, 2017. https://www.sciencedirect.com/science/article/pii/S1877050917317994.

Gonzales, Daniel, and Sarah Harting. *Designing Unmanned Systems with Greater Autonomy*. RR 626-OSD. Santa Monica, CA: RAND, 2014. https://www.rand.org/content/dam/rand/pubs/research_reports/RR600/RR626/RAND_RR626.pdf.

Grant, Rebecca. "Flying Tiger, Hidden Dragon." *Air Force Magazine,* March, 2002. http://www.airforcemag.com/MagazineArchive/Documents/2002/March%202002/0302tiger.pdf.

Hammes, T.X. "The Democratization of Airpower: The Insurgent and the Drone." *War on the Rocks*, October 18, 2016. https://warontherocks.com/2016/10/the-democratization-of-airpower-the-insurgent-and-the-drone/.

Harkins, Gina. "Marines Test New Drone Swarms a Single Operator Can Control." *Military.com*, July 23, 2018. https://www.military.com/defensetech/2018/07/23/marines-test-new-drone-swarms-single-operator-can-control.html.

Hinote, Clint. *Centralized Control and Decentralized Execution: A Catchphrase in Crisis?* RP 2009–1. Maxwell AFB, AL: Air Force Research Institute, 2009. https://permanent.access.gpo.gov/gpo23521/a550460.pdf.

Hoadley, Daniel. *Artificial Intelligence and National Security*. CRS Report No. R45178. Washington Congressional Research Service, 2018. https://fas.org/sgp/crs/natsec/R45178.pdf.

Homans, Charles. "War Games: A Short History." *Foreign Policy,* August 31, 2011. https://foreignpolicy.com/2011/08/31/war-games-a-short-history/.

Honegger, Barbara. "NPS 'Cloud Computing' Lab Up and Running." *Naval Postgraduate School*, February 8, 2010. https://web.nps.edu/About/News/NPS-Cloud-Computing-Lab-Up-and-Running-.html.

Hurst, Jules. "Robotic Swarms in Offensive Maneuver." *Joint Force Quarterly*, no. 87, 2017. http://ndupress.ndu.edu/Publications/Article/1326017/robotic-swarms-in-offensive-maneuver/.

Ilachinski, Andrew. "Artificial Intelligence and Autonomy: Opportunities and Challenges." *Center for Naval Analyses*, 2017. https://www.cna.org/cna_files/pdf/DRM-2017-U-014796-Final.pdf.

Jacobson, Lee. "Applying a Genetic Algorithm to the Traveling Salesman Problem." *Project Spot*, August 20, 2012. http://www.theprojectspot.com/tutorial-post/applying-a-genetic-algorithm-to-the-travelling-salesman-problem/5.

Johnson, Jeffrey S. “Initiative in Soviet Air Force Tactics and Decision Making.” Master’s thesis, Naval Postgraduate School, 1986. https://calhoun.nps.edu/handle/10945/21923.

Kamali, Kaivan, L., Jiang, J. Yen, and K. Wang. “Using Q-Learning and Genetic Algorithms to Improve the Efficiency of Weight Adjustments for Optimal Control and Design Problems.” *Journal of Computing and Information Science in Engineering* 7, no. 4 (December 1, 2007): 302–308. http://computingengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1400904.

Kania, Elsa. “Battlefield Singularity: Artificial Intelligence, Military Revolution, and China’s Future Military Power.” *Center for a New American Security,* November 28, 2017. https://www.cnas.org/publications/reports/battlefield-singularity-artificial-intelligence-military-revolution-and-chinas-future-military-power.

Knight, Will. “The Dark Secret at the Heart of AI.” *MIT Technology Review*. April 11, 2017. https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.

Laslie, Brian D. *The Air Force Way of War: U.S. Tactics and Training after Vietnam*. Lexington, Kentucky: University Press of Kentucky, 2015.

Leombruni, Roberto, and Matteo Richiardi, “Why are Economists Skeptical about Agent-Based Simulations?” *Physica A: Statistical Mechanics and its Applications*, vol. 335, 2005. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.1155&rep=rep1&type=pdf.

Leonard, Matt. “DARPA Looks to Control Drone Swarms with VR.” *Defense Systems*, March 26, 2018. https://defensesystems.com/articles/2018/03/28/darpa-offset-drone-architecture.aspx.

Leoni, Victoria. “New Undersea Drones are Smaller, Cheaper, and Can be Refueled Deep Under Water.” *Defense News*, April 10, 2018. https://www.defensenews.com/digital-show-dailies/navy-league/2018/04/10/new-undersea-drones-are-smaller-cheaper-and-can-be-refueled-deep-under-water/.

Lin, Jeffrey, and P.W. Singer. “China is Making 1,000-UAV Drone Swarms Now.” *Popular Science*, January 8, 2018. https://www.popsci.com/china-drone-swarms.

Lynch, Mike. “AI Cyberattacks Will be Almost Impossible for Humans to Stop.” *Wired*, December 28, 2017. https://www.wired.co.uk/article/ai-cyberattack-mike-lynch/.

Maaranen, Heikki, Kaisa Miettinen, and Antti Penttinen. “On Initial Populations of a Genetic Algorithm for Continuous Optimization Problems.” *Journal of Global Optimization 37*, no. 3 (2007): 405. http://www.cs.uoi.gr/~lagaris/GRAD_GLOPT/projects/genetic_POPULATIONS.pdf.

Maaranen, Heikki, Kaisa Miettinen, and Marko Makela. “Quasi-Random Initial Population for Genetic Algorithms.” *Computers & Mathematics with Applications* 47, no. 12 (2004): 1885–1895. https://core.ac.uk/download/pdf/82606936.pdf.

Majumdar, Dave. “Who Attacked a Russian Military Base with a ‘Swarm’ Strike?” *National Interest*, January 12, 2018. https://nationalinterest.org/feature/who-attacked-russian-military-base-swarm-strike-24060.

Marks, Robert, and Hermann Schnabl. “Genetic Algorithms and Neural Networks: A Comparison Based on the Repeated Prisoners Dilemma.” In *Computational Techniques for Modelling Learning in Economics*, Boston, MA: Springer, 1999. http://www.agsm.edu.au/bobm/papers/jena.pdf.

Mattis, James. “Press Gaggle by Secretary Mattis En Route to Washington, DC.” *Department of Defense Transcripts*, February 17, 2018. https://www.defense.gov/News/Transcripts/Transcript-View/Article/1444921/press-gaggle-by-secretary-mattis-en-route-to-washington-dc/.

Mattis, James. *2018 National Defense Strategy of the United States of America*, Washington, DC: Department of Defense, 2018. https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

McAllister, Brian Linn. “The American Way of War Debate: An Overview.” *Historically Speaking* 11, no. 5, 2010. https://muse.jhu.edu/article/405440/summary.

McCaney, Kevin. “Day of the LOCUST: Navy Demonstrates Swarming UAVs.” *Defense Systems,* April 15, 2015. https://defensesystems.com/articles/2015/04/15/onr-locust-swarming-autonomous-uavs.

McLeary, Paul. “The Pentagon’s Third Offset May Be Dead, But No One Knows What Comes Next.” *Foreign Policy*, December 18, 2017. https://foreignpolicy.com/2017/12/18/the-pentagons-third-offset-may-be-dead-but-no-one-knows-what-comes-next/.

Meden, Alec. “DARPA’s Game of Drones.” *Atlantic Council*, accessed October 9, 2018. http://artoffuturewarfare.org/2016/12/darpas-game-of-drones/.

Mehta, Aaron. “Experiment Over: Pentagon’s Tech Hub Gets a Vote of Confidence.” *Defense News,* August 9, 2018. https://www.defensenews.com/pentagon/2018/08/09/experiment-over-pentagons-tech-hub-gets-a-vote-of-confidence/.

Mehta, Aaron. “National Defense Strategy Released with Clear Priority: Stay Ahead of Russia and China.” *Defense News*, January 19, 2018. https://www.defensenews.com/breaking-news/2018/01/19/national-defense-strategy-released-with-clear-priority-stay-ahead-of-russia-and-china/.

Military Aerospace. “DARPA Adds Two Companies to OFFSET Swarm Reconnaissance Drone Research Project.” May 8, 2018. https://www.militaryaerospace.com/articles/print/volume-29/issue-4/unmanned-vehicles/darpa-adds-two-companies-to-offset-swarm-reconnaissance-drone-research-project.html.

Mittu, Ranjeev, Donald Sofge, Alan Wagner, and William Frere Lawless, *Robust Intelligence and Trust in Autonomous Systems*. Boston, MA: Springer, 2016.

Mnih, Volodymyr, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. Bellemare, A. Graves, M. Riedmiller, A. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. “Human-Level Control through Deep Reinforcement Learning.” *Nature* 518, no. 7540, 2015, https://www.nature.com/articles/nature14236.

Momyer, William W. “The Counter Air Battle (Air Superiority).” In *Airpower in Three Wars [WWII, Korea, Vietnam]*, Maxwell AFB, AL: Air University Press, 2003.

Munoz, Mauricio. “Agent-based Simulation and Analysis of a Defensive UAV Swarm Against an Enemy UAV Swarm.” Master’s thesis, Naval Postgraduate School, 2011. https://calhoun.nps.edu/handle/10945/5700.

Naval Postgraduate School MOVES. “Swarm Commander Tactics.” Accessed December 4, 2017. https://gitlab.nps.edu/moves/swarm-commander-tactics/tree/master.

Nelson, Christopher. “Fleet Tactics Returns – A Conversation with Authors Wayne Hughes and Bob Girrier.” *Center for International Maritime Security*, July 30, 2018. http://cimsec.org/fleet-tactics-returns-a-conversation-with-authors-wayne-hughes-and-bob-girrier/37040.

Newlands, Murray. “The Democratization of Machine Learning is at Hand and this AI Company is on the Front Lines.” *Forbes,* October 20, 2017. https://www.forbes.com/sites/mnewlands/2017/10/20/the-democratization-of-machine-learning-is-at-hand-and-this-ai-company-is-on-the-front-lines/#32586ede2451.

O'Leary, Mary Beth. "Revolutionizing Everyday Products with Artificial Intelligence." *MIT News*, June 1, 2018. http://news.mit.edu/2018/revolutionizing-everyday-products-with-artificial-intelligence-mit-meche-0601.

Osinga, Frans. *Science, Strategy, and War: The Strategic Theory of John Boyd*. London, UK: Routledge, 2007.

Padgett, Nathan. "Defensive Swarm: An Agent-based Modeling Analysis." Master's thesis, Naval Postgraduate School, 2017, https://calhoun.nps.edu/handle/10945/56777.

Parloff, Roger. "Why Deep Learning is Suddenly Changing Your Life." *Fortune*, September 28, 2017. http://fortune.com/ai-artificial-intelligence-deep-machine-learning/.

Peck, Michael. "Why the Pentagon Loves War Games Again." *National Interest*, May 14, 2016. https://nationalinterest.org/feature/why-the-pentagon-loves-war-games-again-16197.

Pellerin, Cheryl. "Project Maven to Deploy Computer Algorithms to War Zone by Year's End." *Department of Defense News*, July 21, 2017. https://www.defense.gov/News/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/.

Porat, Talya, Tal Oron-Gilad, Michal Rottem-Hovev, and Jacob Silbiger. "Supervising and Controlling Unmanned Systems: A Multi-phase Study with Subject Matter Experts." *Frontiers in Psychology* 7, 2016. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00568/full.

Purves, Duncan, Ryan Jenkins, and Bradley Strawser. "Autonomous Machines, Moral Judgment, and Acting for the Right Reasons." *Ethical Theory and Moral Practice* 18, no. 4, 2015. https://www.researchgate.net/publication/276307723_Autonomous_Machines_Moral_Judgment_and_Acting_for_the_Right_Reasons.

Qiwan, Fang, Yin Zhixiang, and Jiang Chuanfu. *Menace of Anti-Ship Missiles and Shipborne Laser Weapons*. NAIC-ID(RS)T0337-96. Wright-Patterson AFB, OH: National Air Intelligence Center, 1996. http://www.dtic.mil/dtic/tr/fulltext/u2/a313312.pdf.

RAND. "Wargaming." Accessed September 20, 2018. https://www.rand.org/topics/wargaming.html.

Rassler, Don. "Remotely Piloted Innovation: Terrorism, Drones, and Supportive Technology." *Combating Terrorism Center*, October 20, 2016. https://ctc.usma.edu/remotely-piloted-innovation-terrorism-drones-and-supportive-technology/.

Retter, Lucia, Alexandra Hall, James Black, Nathan Ryan. *Moral Component of Cross-Domain Conflict*. RR 1505-MOD. Santa Monica, CA: RAND, 2016. https://www.rand.org/pubs/research_reports/RR1505.html.

Richter, James. “On Mutation and Crossover in the Theory of Evolutionary Algorithms.” PhD diss., Montana State University, 2010. https://www.cs.montana.edu/techreports/0910/Richter.pdf.

Robbins, Jim. “America’s Red Army.” *New York Times*, April 17, 1988. https://www.nytimes.com/1988/04/17/magazine/americ-s-red-army.html.

Rodriguez, Jesus. “Understanding Hyperparameters Optimization in Deep Learning Models: Concepts and Tools.” *Medium*, August 8, 2018. https://towardsdatascience.com/understanding-hyperparameters-optimization-in-deep-learning-models-concepts-and-tools-357002a3338a.

Rosenblatt, Frank. “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.” *Psychological Review* 65, no. 6, 1958. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.3398&rep=rep1&type=pdf.

Rosenthal, Daniel, and Loren Schulman. “Trump’s Secret War on Terror.” *Atlantic*, August 10, 2018. https://www.theatlantic.com/international/archive/2018/08/trump-war-terror-drones/567218/.

Ryan McCune, R. Purta, M. Dobski, A. Jaworski, G. Madey, A. Madey, Y. Wei, M. Blake. “Investigations of DDDAS for Command and Control of UAV Swarms with Agent-Based Modeling.” In *Proceedings of the 2013 Winter Simulation Conference: Making Decisions in a Complex World* (December 2013), 1467–1478. https://ieeexplore.ieee.org/document/6721531.

Sabin, Philip. *Simulating War: Studying Conflict through Simulation Games.* London, UK: Continuum International Publishing Group, 2014.

Sahin, Erol. *Swarm Robotics: From Sources of Inspiration to Domains of Application*. METU-CENG-TR-2005-01. Ankara, Turkey: Middle East Technical University, 2005. http://www.kovan.ceng.metu.edu.tr/pub/pdf/METU-CENG-TR-2005-01.pdf.

Sanchez, Raf. “Russia uses Missiles and Cyber Warfare to fight off ‘Swarm of Drones’ Attacking Military Bases in Syria.” *Telegraph*, January 9, 2018. https://www.telegraph.co.uk/news/2018/01/09/russia-fought-swarm-drones-attacking-military-bases-syria/.

Sanders, Andrew. “Drone Swarms.” Monograph, United States Army Command and General Staff College, 2017. http://www.dtic.mil/docs/citations/AD1039921.

Savitz, Scott, I. Blickstein, P. Buryk, R. Button, P. DeLuca, J. Dryden, J. Mastbaum, J. Osburg, P. Padilla, A. Potter, C. Price, L. Thrall, S. Woodward, R. Yardley, J. Yurchak. *U.S. Navy Employment Options for Unmanned Surface Vehicles*. RR-384-NAVY. Santa Monica, CA: RAND, 2013. https://www.rand.org/content/dam/rand/pubs/research_reports/RR300/RR384/RAND_RR384.pdf.

Schehl, Matthew, and Khaboshi Imbukwa. "Student Wargaming Activities Address Sponsors' Direct Needs." *Naval Postgraduate School*, July 11, 2018. https://my.nps.edu/-/student-wargaming-activities-address-sponsors-direct-needs.

Schehl, Matthew. "JIFX Continues to Help DoD, Academia Explore Limits of New Technology." August 9, 2018. https://www.dvidshub.net/news/295041/jifx-continues-help-dod-academia-explore-limits-new-technology.

Schwartz, Howard M. "Multi-agent Machine Learning: A Reinforcement Approach." Hoboken, NJ: John Wiley & Sons, 2014. https://ebookcentral.proquest.com/lib/ebook-nps/detail.action?docID=1775207.

Segarra, Lisa. "This Racing Drone Just Set a Guinness World Speed Record." *Fortune,* July 14, 2017. http://fortune.com/2017/07/14/fastest-drone-guinness-world-record/.

Serafino, Loris. *Between Theory and Practice: Guidelines for an Optimization Scheme with Genetic Algorithms—Part I*. Shenzhen, China: Kuang-Chi Institute of Advanced Technology, 2011. https://arxiv.org/pdf/1112.4323.pdf.

Shane, Scott, and Daisuke Wakabayashi. "'The Business of War': Google Employees Protest Work for the Pentagon." *New York Times*, April 4, 2018. https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

Shlapak, David. "The Russian Challenge." PE 250-A. Santa Monica, CA: RAND, 2018. https://www.rand.org/pubs/perspectives/PE250.html.

Silver, David, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis. "Mastering the Game of Go Without Human Knowledge." *Nature* 550, no. 7676, 2017. https://deepmind.com/research/publications/mastering-game-go-without-human-knowledge/.

Singer, Peter W. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. New York, NY: Penguin Press, 2009.

Smith, Michael. “Kinetic Strike for Special Tactics to Achieve Precision Strike Effects.” Presentation at U.S. Air Force Weapons School, Nellis AFB, NV, 2016. https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2016/psts/CaptMichaelDSmith.pdf.

Smith, Roger. “The Long History of Gaming in Military Training.” *Simulation & Gaming* no. 41, 2010. http://journals.sagepub.com.libproxy.nps.edu/doi/pdf/10.1177/1046878109334330.

Storr, Jim. “A Command Philosophy for the Information Age: The Continuing Relevance of Mission Command.” *Defense Studies* 3, no. 3, 2003. https://www.tandfonline.com/doi/pdf/10.1080/14702430308405081.

Straffin, Philip D. *Game Theory and Strategy*. Washington, DC: Mathematical Association of America, 1993.

Sultan, Nabil. “Cloud Computing: A Democratizing Force?” *International Journal of Information Management* 33, no. 5 (October 2013): 810-815. https://www.sciencedirect.com/science/article/pii/S0268401213000820.

Sutton, Richard, and Andrew Barto. *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press, 2014. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf.

Thomas, William. “Meta-Calculations and the Mathematics of War.” In *Rational Action: The Sciences of Policy in Britain and America, 1940–1960,* edited by Jed Buchwald, 99–102. Cambridge, MA: MIT Press, 2015.

Toft, Monica. “The Dangerous Rise of Kinetic Diplomacy.” *War on the Rocks*, May 14, 2018. https://warontherocks.com/2018/05/the-dangerous-rise-of-kinetic-diplomacy/.

Trevithick, Joseph. “Navy’s Sea Hunter Drone Ship Is Getting a New Owner, New Abilities, and a Sister.” *Drive*, February 6, 2018. http://www.thedrive.com/the-war-zone/18264/navys-sea-hunter-drone-ship-is-getting-a-new-owner-new-abilities-and-a-sister.

United States Air Force Office of the Chief Scientist. *Autonomous Horizons: System Autonomy in the Air Force—Volume I: Human-Autonomy Teaming*. Langley AFB, VA: Headquarters of the Air Force, 2015. https://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf.

Vincent, James. “AI Bots Trained for 180 Years a Day to Beat Humans at DOTA 2.” *Verge*, June 25, 2018. https://www.theverge.com/2018/6/25/17492918/openai-dota-2-bot-ai-five-5v5-matches.

Voss, Peter. “From Narrow to General AI.” *Medium*, October 3, 2017. https://medium.com/intuitionmachine/from-narrow-to-general-ai-e21b568155b9.

Wakabayashi, Daisuke, and Scott Shane. “Google Will Not Renew Pentagon Contract That Upset Employees.” *New York Times*, June 1, 2018. https://www.nytimes.com/2018/06/01/technology/google-pentagon-project-maven.html.

Wong, Yuna Huh. “How Can Gaming Help Test Your Theory?” *RAND Blog*, May 18, 2016. https://www.rand.org/blog/2016/05/how-can-gaming-help-test-your-theory.html.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California