

T.C.
MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKA ANABİLİM DALI

DAHA GERÇEKÇİ GERÇEK ZAMANLI STRATEJİ OYUNLARINA DOĞRU:
SİS VE ŞANS FAKTÖRLERİNİN ETKİSİNİN MODELLENMESİ

YÜKSEK LİSANS BİLİM TEZİ

YAKUP ÖNÜR

ARALIK 2024
MUĞLA

MUGLA SITKI KOÇMAN ÜNİVERSİTESİ

Fen Bilimleri Enstitüsü

TEZ ONAYI

YAKUP ÖNÜR tarafından hazırlanan **DAHA GERÇEKÇİ GERÇEK ZAMANLI STRATEJİ OYUNLARINA DOĞRU: SİS VE ŞANS FAKTÖRLERİNİN ETKİSİNİN MODELLENMESİ** başlıklı tezinin, 24/12/2024 tarihinde aşağıdaki jüri tarafından Yapay Zeka Anabilim Dalı'nda yüksek lisans derecesi için gerekli şartları sağladığı oybirliği/oyçokluğu ile kabul edilmiştir.

TEZ SINAV JURİSİ

Dr.Öğr.Üyesi Enis KARAARSLAN (**Jüri Başkanı**)(**Danışman**) İmza:
Bilgisayar Mühendisliği Bölümü / Siber Güvenlik ABD.
Muğla Sıtkı Koçman Üniversitesi, Muğla

Dr.Öğr.Üyesi Nida GÖKÇE NARİN (**Üye**) İmza:
İstatistik Bölümü
Muğla Sıtkı Koçman Üniversitesi, Muğla

Dr.Öğr.Üyesi Feriştah DALKILINÇ (**Üye**) İmza:
Bilgisayar Mühendisliği Bölümü
Dokuz Eylül Üniversitesi, İzmir

ANA BİLİM DALI BAŞKANLIĞI ONAYI

Dr.Öğr.Üyesi Nida GÖKÇE NARİN İmza:
Yapay Zeka Ana Bilim Dalı Başkanı
Muğla Sıtkı Koçman Üniversitesi, Muğla

Dr.Öğr.Üyesi Enis KARAARSLAN İmza:
Danışman, Bilgisayar Mühendisliği Bölümü / Siber Güvenlik ABD.
Muğla Sıtkı Koçman Üniversitesi, Muğla

Savunma Tarihi: 24/12/2024

Tez çalışmalarım sırasında elde ettiğim ve sunduğum tüm sonuç, doküman, bilgi ve belgelerin tarafımdan bizzat ve bu tez çalışması kapsamında elde edildiğini; akademik ve bilimsel etik kurallarına uygun olduğunu beyan ederim. Ayrıca, akademik ve bilimsel etik kuralları gereği bu tez çalışması sırasında elde edilmemiş başkalarına ait tüm orijinal bilgi ve sonuçlara atıf yapıldığını da beyan ederim.

Yakup Önür
24/12/2024

ÖZET

DAHA GERÇEKÇİ GERÇEK ZAMANLI STRATEJİ OYUNLARINA DOĞRU: SİS VE ŞANS FAKTÖRLERİNİN ETKİSİNİN MODELLENMESİ

Yakup Önür

Yüksek Lisans Tezi

Fen Bilimleri Enstitüsü

Yapay Zeka Anabilim Dalı

Danışman: Dr.Öğr.Üyesi Enis KARAARSLAN

Aralık 2024, 57 sayfa

Simülasyon temelli eğitim, belirli senaryoların tekrarıyla personel becerilerini geliştirmekte yaygınlaşmaktadır. Gerçek Zamanlı Strateji (RTS) oyunları, askeri eğitim veya stratejilerin test edilmesinde önemli bir potansiyel sunmaktadır. Bu çalışma, yapay zeka kullanarak RTS oyunlarında gerçekçiliği artırmayı, özellikle Capture the Flag (CTF) çerçevesinde “savaş sisini” (fog of war) ve şans faktörlerini entegre etmeyi amaçlamaktadır. Savaş sisi, yapay zeka ajanlarının kısıtlı bilgiyle karar vermesini sağlayarak belirsizliği artırır ve daha dinamik stratejik düşünmeyi zorunlu kılar. Araştırma sorusu, “YZ destekli RTS oyunlarında sis ve şans faktörleri stratejik simülasyonların gerçekçiliğini ve stratejik düşünme yeteneklerini nasıl etkiler?” olarak ele alınmıştır. Bu amaçla, “Gym-Cap” platformunda sis ve şans faktörleri test edilmiş, platformun kısıtları nedeniyle tarafımızca temel düzeyde bir CTF oyunu kodlanmış ve üzerinde Q-Öğrenme, Gelişmiş Q-Öğrenme, Çift Q-Öğrenme, SARSA, Beklenen SARSA, Monte Carlo ve Dyna-Q algoritmaları karşılaştırılmıştır. Sonuçlar, sis ve şans faktörlerinin belirsizliği artırarak stratejik karar alma süreçlerini zenginleştirdiğini göstermiştir. Yapay zeka ajanları; askeri eğitim, siber güvenlik ve afet yönetimi gibi alanlarda daha gerçekçi senaryolar sunabilir. Gelecekte daha güçlü donanımlar ve karmaşıklaştırılmış senaryolar ve parametreler kullanılarak takım koordinasyonu, iletişim kısıtları ve oyun teorisi ilkeleriyle stratejik analizler derinleştirilebilir.

Anahtar kelimeler: Yapay Zeka, Gerçek Zamanlı Strateji Oyunları, Pekiştirmeli Öğrenme, Simülasyon Tabanlı Eğitim, Bayrağı Yakala

ABSTRACT

TOWARDS MORE REALISTIC REAL-TIME STRATEGY GAMES: MODELING THE EFFECT OF FOG AND CHANCE FACTORS

Yakup ÖNÜR

Master of Science (M.Sc.)

Graduate School of Natural and Applied Sciences

Department of Artificial Intelligence

Supervisor: Asst.Prof.Dr. Enis KARAARSLAN

November 2024, 57 pages

Simulation-based training enhances skills by repeating specific scenarios. Real-Time Strategy (RTS) games provide significant potential for military training or strategy testing without real conflict. This study aims to enhance RTS game realism using artificial intelligence (AI) by integrating fog of war and chance factors within a Capture the Flag (CTF) framework. The fog of war forces AI agents to operate with limited information, increasing uncertainty and fostering dynamic strategic reasoning. The research question addressed is: “How do fog and chance factors in AI-supported RTS games affect the realism of strategic simulations and strategic thinking abilities?” To explore this, fog and chance were tested on the “Gym-Cap” platform. Due to its limitations, a simplified CTF environment was coded, where reinforcement learning algorithms, including Q-Learning, Improved Q-Learning, Double Q-Learning, SARSA, Expected SARSA, Monte Carlo, and Dyna-Q, were compared. Results show that integrating fog and chance enriches strategic decision-making by increasing uncertainty. AI agents can create more realistic scenarios applicable to military training, cybersecurity, and disaster management. Future studies can leverage more powerful hardware and complex CTF environments to explore team coordination, communication constraints, and game theory for deeper strategic analysis.

Keywords: Artificial Intelligence, Real-Time Strategy Games, Reinforcement Learning, Simulation-Based Training, Capture the Flag

ÖNSÖZ

Bu çalışmayı hazırlarken bana her zaman destek olan aileme, motivasyonunu hiç eksik etmeyen değerli arkadaşlarıma ve akademik yolculuğum boyunca rehberliğini ve bilgeliğini esirgemeyen danışman hocam Dr. Öğr. Üyesi Enis Karaarslan'a sonsuz teşekkürlerimi sunarım. Onların desteği olmadan bu çalışmayı tamamlamak mümkün olmazdı. Hepinizin bana kattıkları için minnettarım.



İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER	viii
ÇİZELGE LİSTESİ	ix
ŞEKİL LİSTESİ	x
1. GİRİŞ	1
2. TEMELLER	4
2.1. Strateji Oyunları	4
2.2. Strateji Oyunlarında Yapay Zeka	5
2.3. Yapay Zeka Yöntemlerinin Temelleri	6
2.3.1. Pekiştirmeli öğrenme (Reinforcement Learning)	6
2.3.2. Q-Öğrenme (Q-Learning)	6
2.3.3. Gelişmiş Q-Öğrenme (Developed Q-Learning)	7
2.3.4. SARSA	7
2.3.5. Çift Q-Öğrenme (Double Q-Learning)	8
2.3.6. Beklenen SARSA (Expected SARSA)	8
2.3.7. Monte Carlo	9
2.3.8. Dyna-Q	9
3. LİTERATÜR ÇALIŞMASI	11
3.1. Strateji Oyunları	11
3.2. Pekiştirmeli Öğrenme	12
4. YÖNTEM / SİSTEM ÖNERİSİ	15
4.1. CTF Test Ortamı	16
4.2. Öğrenme Metodolojisi: Yapay Zeka Ajanları	18
4.3. Test Metodolojisi	20
4.3.1. Gelişmiş test yöntemi	24
4.3.2. Sis ve şans faktörleri	25
4.4. Temel Düzey CTF	27
5. UYGULAMA	29

5.1. Gym-Cap Uygulaması	29
5.2. Temel Düzey CTF Oyunu	32
5.2.1. Deneysel kurulum ve sonuçlar	33
5.2.2. Şans ve sis faktörlerinin entegrasyonu	33
6. SONUÇLAR	35
6.1. Gym-Cap Optimizasyon Entegrasyonu	36
6.2. Gym-Cap Sonuçları	37
6.3. Gym-Cap Rambo Sonuçları	40
6.4. Pekiştirmeli Modellerin Karşılaştırma Sonuçları	41
6.5. Q-öğrenme Model Karşılaştırma Sonuçları	43
6.6. Şans Sonuçları	45
6.7. Sis Sonuçları	46
7. TARTIŞMA	48
7.1. Araştırma Sorusu Üzerine Bulgular	48
7.2. Çalışmanın Sınırlamaları	49
7.3. Gelecek Araştırmalar	50
8. SONUÇ	52
KAYNAKÇA	54

ÇİZELGE LİSTESİ

Çizelge 5.1. Gerçekçiliği ve Karmaşıklığı Artırmak İçin Simülasyon Parametre Ayarlamaları	31
Çizelge 6.2. YZ Savunma ve Saldırgan Ajanlarının Stratejileri	38
Çizelge 6.3. Savunma ve Saldırgan Ajanlarının Çeşitli Saldırı Stratejilerine Karşı Ortalama Puanları (Bölüm 1)	38
Çizelge 6.4. Savunma ve Saldırgan Ajanlarının Çeşitli Saldırı Stratejilerine Karşı Ortalama Puanları (Bölüm 2)	38
Çizelge 6.5. Roomba Stratejisi ile Mavi Takım Stratejileri Arasındaki Puanlar	40

ŞEKİL LİSTESİ

Şekil 4.1.	CTF Simülasyon Çerçevesi Akış Diyagramı	17
Şekil 4.2.	CTF Gym Tam Ekran. (a) Tüm oyuncular ve bayraklar. (b) Sadece bayraklar. (c)-(d) Takım görünümleri. (e)-(f) Takım alanları. (g)-(h) Bireysel alanlar.	17
Şekil 4.3.	Çerçevede Aktörler, Varlıklar ve Faktörler	21
Şekil 6.4.	YZ Ajanının Öğrenme İlerlemesi: PPO Eğitimi Sırasında Her Değerlendirme Aralığındaki Ortalama Ödül	37
Şekil 6.5.	Çeşitli Saldırı Türlerine Karşı Savunma Stratejileri İçin Ortalama Puanlar	39
Şekil 6.6.	Ortalama Puanlar Üzerindeki Şans ve Sis Koşullarının Etkisi .	39
Şekil 6.7.	Farklı Koşullar Altında Roomba Stratejisi İçin Ortalama Puanlar	41
Şekil 6.8.	Farklı Modellerin Sonuçlarının Karşılaştırması	42
Şekil 6.9.	Farklı Modellerin CPU-RAM Karşılaştırması	42
Şekil 6.10.	Farklı Modellerin Zaman ve Adım Karşılaştırması	43
Şekil 6.11.	QL-Normal ve QL-Developed Modellerinin Performans Karşılaştırması	44
Şekil 6.12.	QL-Normal ve Şans Faktörleri İçeren Modellerin Performans Karşılaştırması	46
Şekil 6.13.	QL-Normal ve Sis Faktörleri İçeren Modellerin Performans Karşılaştırması	47

SEMBOLLER VE KISALTMALAR

CPU	Central Processing Unit
CTF	Capture the Flag
GPU	Graphics Processing Unit
Gym-Cap	OpenAI Gym tabanlı Capture the Flag ortamı
HRL	Hiyerarşik Pekiştirmeli Öğrenme (Hierarchical Reinforcement Learning)
MCTS	Monte Carlo Tree Search (Monte Carlo Ağaç Arama)
MlpPolicy	Multi-Layer Perceptron Policy
MOBA	Multiplayer Online Battle Arena
PPO	Proximal Policy Optimization
RAM	Random Access Memory
RL	Pekiştirmeli Öğrenme (Reinforcement Learning)
RTS	Real-Time Strategy (Gerçek Zamanlı Strateji)
SARSA	State-Action-Reward-State-Action
SQL	Structured Query Language
YZ	Yapay Zeka (Artificial Intelligence)

1. GİRİŞ

Gerçek Zamanlı Strateji (RTS) oyunları, oyuncuların anında karar alıp durumlara uyum sağlamasını gerektiren oyunlardır. Bu oyunlar, askeri çatışmaları veya strateji simülasyonlarını taklit ederek eğitim, analiz ve eğlence amaçları için kullanılmaktadır. RTS oyunları, oyunculara karmaşık ve belirsiz durumlarla başa çıkma becerilerini geliştirme fırsatı sunar, böylece stratejik düşünme yeteneklerini artırır. Bu oyunlardaki temel kavramlardan biri olan 'savaş sis' (fog of war), oyuncuların savaş alanındaki sınırlı görüş ve belirsizlik deneyimini temsil eder. Bu mekanik, gerçek dünya koşullarını simüle ederken, oyuncuların düşmanın hareketleri hakkında tam bilgiye sahip olmadan stratejik kararlar almasını zorunlu kılar. Savaş sisinin simülasyonu, savaş oyunlarının gerçekçiliğini artırmak için esastır. Obeia ve Wójcik (2023), dijital savaş oyunlarındaki farklı 'savaş sis' modellerini karşılaştırmış ve stratejik simülasyonlarda etkili sis modellerinin önemini vurgulamıştır. Ayrıca, oyunlar ve simülatörler, komuta ve kontrol uygulamalarında yapay zekanın geliştirilmesi için değerli platformlar olarak giderek daha fazla tanınmaktadır (2023). Ancak, bu simülasyonların potansiyelini tam anlamıyla gerçekleştirebilmek için gerçekçiliğin artırılması gerekmektedir. Bu bağlamda, oyunlara şans (rastgele olaylar) ve belirsizlik (savaş sis etkileri) unsurlarının entegre edilmesi, oyun deneyimini daha gerçekçi ve eğitici hale getirmektedir.

Bu çalışma, yapay zeka kullanarak RTS oyunlarında gerçekçiliği artırmayı, özellikle Bayrağı Yakalama (Capture the Flag CTF) çerçevesinde "savaş sisini" entegre etmeyi amaçlamaktadır. Bu kavram, oyuncuların veya yapay zeka ajanlarının savaş alanında kısıtlı bilgiyle hareket etmesini temsil eder. Belirsizliğin bu şekilde modellenmesi, gerçek dünya koşullarını daha yakından taklit ederek stratejik karar

almayı güçleştirmekte, öngörülemezlik unsuru ekleyerek oyunun simülasyon değerini artırmaktadır. Yapay zeka modelleri, dinamik savaş sis ve şans faktörlerine uyum sağlamak için sezgisel öğrenme yöntemleri kullanılarak test edilmiştir. Bu unsurlar, yapay zekanın eksik bilgiyle karar verme süreçlerini taklit etmesini sağlayarak, gerçek hayattaki karmaşıklığı eklemektedir. Bulgular, savaş sisinin ve şans faktörlerinin entegre edilmesinin oyun sonuçları üzerinde önemli etkileri olduğunu göstermektedir; bu da belirsizliğin stratejik karar alma süreçlerindeki rolünü vurgulamaktadır. Oyuncular, kısıtlı bilgi ve şans olaylarına dayalı stratejilerini yeniden uyarlamak zorunda kaldıklarında daha gerçekçi davranışlar sergilemektedir. Bu ilerleme, askeri eğitim için faydalı olup, belirsizlik altında karar verme becerilerini ve uyum sağlama yeteneklerini geliştirmektedir. Bu çalışma, yapay zekanın oyun gerçekçiliğini nasıl artırabileceğini göstererek, stratejik savaş oyunlarının askeri eğitim ve planlama araçları olarak gücünü pekiştirmektedir.

Araştırma sorumuz, Yapay zeka destekli RTS oyunlarında sis (belirsizlik ve bilgi kıtlığı) ve şans (rastgelelik ve şans) faktörlerinin entegrasyonu, stratejik simülasyonların gerçekçiliğini ve oyuncuların stratejik düşünme yeteneklerini nasıl etkiler? şeklindedir.

Bu çalışmanın geri kalanı şu şekilde düzenlenmiştir: Bölüm 2, strateji oyunlarının temellerini, yapay zekanın stratejik kararlara etkisini ve bu alandaki temel kavramları sunmaktadır. Bölüm 3, sis ve şans faktörlerinin stratejik simülasyonlardaki rolünü irdeleyen ilgili literatürü gözden geçirmektedir. Bölüm 4, CTF çerçevesinde sis ve şans faktörlerinin entegrasyonunu sağlayan yöntem ve sistem önerisini aktarmaktadır. Bölüm 5, geliştirilen uygulamayı tanıtmakta, Gym-Cap ve temel düzey CTF ortamlarında yapılan deneylerin ayrıntılarını paylaşmaktadır. Bölüm 6, elde edilen sonuçları sunmakta ve sis ile şansın stratejik

karar alma süreçleri üzerindeki etkilerini analiz etmektedir. Bölüm 7, tartışma bölümünde bulgularımızı değerlendirip kısıtları ve geleceğe yönelik çalışmaları ele alırken; Bölüm 8, genel bir sonuç ve değerlendirme yaparak çalışmayı sonlandırmaktadır.



2. TEMELLER

2.1. Strateji Oyunları

RTS oyunları, oyuncuların anında karar alıp uygulaması gereken oyunlardır. Örneğin, asker ordusunu yönetirken, oyuncu aynı zamanda kaynak yönetimi ve ekonomi konularını da dikkate almalıdır. Popüler RTS oyunları arasında "StarCraft" ve "Age of Empires" gibi oyunlar bulunmaktadır (2022).

Savaş oyunu (wargame), gerçek veya senaryo bazlı askeri operasyonların simülasyonu olarak tanımlanabilir. Bu oyunlar genellikle gerçek veya tarihi askeri operasyon senaryolarını kullanır ve oyuncuların birimleri komuta edip yönlendirmesine olanak tanır. Bu oyunlar genellikle savaş stratejisi, taktikler veya lojistik gibi konuları içerir. Savaş oyunları genellikle PC, konsol veya mobil platformlarda oynanabilir. Popüler savaş oyunları arasında "Warhammer 40,000: Dawn of War," "Panzer General" ve "Steel Panthers" gibi oyunlar bulunmaktadır (2022).

Çok Oyunculu Çevrimiçi Savaş Arenası (MOBA) oyunları, strateji oyunlarının bir alt türüdür ve takımların benzersiz yeteneklere sahip bireysel karakterleri kontrol ettiği, birbirlerine karşı mücadele ettiği oyunlardır. Popüler MOBA oyunları arasında "League of Legends" ve "Dota 2" gibi oyunlar yer almaktadır. Bu oyunlar, takım çalışması, stratejik planlama ve gerçek zamanlı karar alma gerektirir, oyuncuların takım arkadaşlarıyla koordineli olarak hareket etmesini ve oyun içi dinamik durumlardaki değişikliklere uyum sağlamasını zorunlu kılar.

CTF oyunları, her takımın kendi bayrağına sahip olduğu ve diğer takımın bayrağını ele geçirme amacını taşıdığı oyunlardır. CTF oyunları çeşitli türlerde olabilir, ancak

en yaygın olanları fiziksel veya sanal ortamlarda gerçekleşen oyunlardır ya da güvenlik yarışmaları şeklinde düzenlenir. Güvenlik yarışmalarında, oyuncular genellikle bilgisayar sistemlerini veya ağlarını ele geçirme tekniklerini öğrenir ve uygularlar. DEFCON CTF gibi yarışmalar buna örnek teşkil etmektedir (2022). CTF oyunları ayrıca, bilgisayar güvenliği ve BT adli bilişim derslerinde öğrenmeyi artırmak için yaygın olarak kullanılmaktadır (2022; 2023). Açık kaynaklı CTF ortamları, eğitim ve araştırma amaçları için faydalı çeşitli işlevsellikler ve yapılandırmalar sunmaktadır (2020).

2.2. Strateji Oyunlarında Yapay Zeka

Yapay Zeka (YZ), strateji oyunlarının vazgeçilmez bir bileşeni haline gelmiş olup, öğrenme, akıl yürütme ve problem çözme gibi insan bilişsel yeteneklerini kullanarak oyun oynama karmaşıklığını ve etkileşimini artırmaktadır. Yapay zekanın oyun içi karar alma süreçlerini otomatikleştirmesi, oyunculara rekabet veya işbirliği için zeki varlıklar sunmakta, bu da gerçek zamanlı strateji oyunlarında özellikle belirgin hale gelmektedir. Bu gelişme, sadece YZ ile daha insana benzer etkileşimler sağlamamakta, aynı zamanda oyun tasarımcılarına oyun mekaniklerini geliştirme, oyuncu davranışlarını analiz etme ve oyun deneyimini bireysel tercihlere göre özelleştirme konusunda paha biçilmez fırsatlar sunmaktadır.

YZ teknolojileri arasında derin öğrenme, sinir ağlarına dayalı olarak, oyuncu davranışlarını taklit etme ve karmaşık stratejiler geliştirme konusunda öne çıkmaktadır. Aynı zamanda, pekiştirmeli öğrenme, yapay zekanın belirlenen hedeflere ulaşmak için en etkili eylemleri belirlemesine olanak tanırken, Monte Carlo Ağaç Arama (MCTS) potansiyel hamleleri ve sonuçlarını yapılandırılmış bir ağaç çerçevesinde keşfetmeyi desteklemektedir. Bu yöntemler, yapay zekanın

karmaşık karar matrislerini benimsemesine ve uygulamasına olanak tanımaktadır.

2.3. Yapay Zeka Yöntemlerinin Temelleri

Bu çalışmada, farklı yapay zeka algoritmaları kullanılarak oyun ortamında ajanların performansı değerlendirilmiştir. Bu bölümde, kullanılan her bir modelin çalışma prensipleri ve uygulama bağlamındaki rolleri detaylı olarak incelenmiştir.

2.3.1. Pekiştirmeli öğrenme (Reinforcement Learning)

Pekiştirmeli öğrenme, bir ajanın bir çevre ile etkileşim kurarak eylemlerini optimize etmeye çalıştığı bir öğrenme paradigmasıdır. Ajan, her eylemden sonra bir ödül veya ceza alır ve amacı, toplam ödülü maksimize eden bir politika geliştirmektir. Temel bileşenleri şunlardır:

- **Ajan (Agent):** Çevre ile etkileşime giren ve eylem seçen varlık.
- **Çevre (Environment):** Ajanın etkileşimde bulunduğu ve geri bildirim aldığı sistem.
- **Politika (Policy):** Ajanın belirli bir durumda hangi eylemi seçeceğini belirleyen strateji.
- **Ödül (Reward):** Ajanın eylemlerinin sonucunda aldığı geri bildirim.
- **Durum (State):** Ajanın çevredeki mevcut durumu.

2.3.2. Q-Öğrenme (Q-Learning)

Q-Öğrenme algoritması, bir ajan için durum-eylem çiftlerine ilişkin $Q(s, a)$ değerlerini öğrenerek gelecekteki ödülleri tahmin eder. Ajan, keşif (ϵ) ve sömürü

arasında denge sağlamak amacıyla ϵ -greedy stratejisi kullanır. Bu stratejide, belirli bir olasılık (ϵ) ile rastgele eylemler seçilerek keşif yapılırken, geri kalan durumda en yüksek Q -değerine sahip eylemler tercih edilerek sömürü sağlanır. Q -değerleri, Formül 2.1'de gösterildiği şekilde güncellenir:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \cdot \left[r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a) \right] \quad (2.1)$$

Bu çalışmada, Q -Öğrenme algoritmasına otomatik parametre optimizasyonu eklenmiştir. Öğrenme hızı (α) ve keşif oranı (ϵ), her bölümde dinamik olarak güncellenerek ajanın başlangıçta çevreyi keşfetmesini, daha sonra ise sömürüye odaklanarak daha iyi stratejiler geliştirmesini sağlamıştır.

2.3.3. Gelişmiş Q -Öğrenme (Developed Q-Learning)

Gelişmiş Q -Öğrenme modeli, temel Q -Öğrenme algoritmasının sabit parametrelerle çalışan bir varyasyonudur. Bu modelde, otomatik parametre optimizasyonu bulunmamaktadır. Sabit öğrenme hızı ($\alpha = 0.2$) ve sabit keşif oranı ($\epsilon = 0.2$) kullanılarak, ajanın daha istikrarlı bir öğrenme süreci geçirmesi hedeflenmiştir. Ancak sabit parametreler, keşif yeteneğini kısıtlayarak çevresel değişimlere daha az adaptasyon yeteneği sunmaktadır. Bu model de Q -değerlerini Formül 2.1'de gösterilen Q -Öğrenme güncellemesi ile hesaplar, ancak sabit parametreler kullanır.

2.3.4. SARSA

SARSA (*State-Action-Reward-State-Action*) algoritması, Q -Öğrenme'ye benzerlik gösterir; ancak Q -değerlerini güncellerken bir sonraki eylemi (a') de hesaba katar. Bu sayede, ajan durumlara özgü ödülleri daha hassas bir şekilde öğrenir. SARSA

güncelleme kuralı, Formül 2.2’de gösterilmiştir:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \cdot [r + \gamma \cdot Q(s', a') - Q(s, a)] \quad (2.2)$$

SARSA, gelecekteki eylemlerin doğrudan etkilerini değerlendirdiği için daha “yumuşak” bir öğrenme stratejisi benimser ve çevreye daha hassas bir adaptasyon sağlar.

2.3.5. Çift Q-Öğrenme (Double Q-Learning)

Çift Q-Öğrenme, Q-Öğrenme’nin ödül tahminlerindeki aşırı optimizmi gidermek amacıyla geliştirilmiştir. İki ayrı Q-tablosu (Q_1 ve Q_2) kullanır ve her güncellemede bu tablolardan biri rastgele seçilerek güncelleme yapılır. Böylece, öğrenme sürecinde dengeli ve daha güvenilir ödül tahminleri elde edilir. Çift Q-Öğrenme güncellemesi, Formül 2.3’de verildiği şekildedir:

$$Q_1(s, a) \leftarrow Q_1(s, a) + \alpha \cdot \left[r + \gamma \cdot Q_2 \left(s', \arg \max_{a'} Q_1(s', a') \right) - Q_1(s, a) \right] \quad (2.3)$$

Bu yaklaşım, Q-değerlerindeki aşırı iyimserliği azaltarak daha kararlı bir öğrenme süreci sunar.

2.3.6. Beklenen SARSA (Expected SARSA)

Beklenen SARSA algoritması, durumdaki tüm olası eylemleri ve bu eylemlerin ortaya çıkma olasılıklarını göz önünde bulundurarak güncelleme yapar. Bu sayede, beklenen ödülleri hesaba katar ve daha az değişken bir öğrenme süreci elde eder. Beklenen

SARSA güncelleme kuralı, Formül 2.4'de sunulmuştur:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \cdot \left[r + \gamma \cdot \sum_{a'} \pi(a'|s')Q(s', a') - Q(s, a) \right] \quad (2.4)$$

Bu yaklaşım, belirsizlik durumlarında daha dengeli performans sergiler.

2.3.7. Monte Carlo

Monte Carlo algoritması, tam bir bölümün tamamlanmasının ardından elde edilen ödülleri kullanarak Q-değerlerini günceller. Her bir durum-eylem çifti için, ilgili gelecek ödüllerin toplamı, Formül 2.5'de gösterildiği gibi hesaplanır:

$$G_t = \sum_{k=0}^{T-t} \gamma^k r_{t+k} \quad (2.5)$$

Bu değerlere dayalı olarak Q-değerlerini günceller. Bölüm bazında öğrenme sağladığı için, geçmiş deneyimlerin daha geniş kapsamlı bir değerlendirmesi yapılır ve dinamik ortamlarda etkili bir yöntemdir.

2.3.8. Dyna-Q

Dyna-Q, model tabanlı pekiştirmeli öğrenme algoritmaları arasında yer alır. Ajan, çevreden topladığı deneyimlerden bir çevre modeli oluşturur. Bu modelden hareketle sanal adımlar simüle edilir ve planlama yapılır. Model güncellemeleri, Formül 2.6'de gösterilmektedir:

$$M(s, a) = (r, s') \quad (2.6)$$

Bu modelden yararlanarak elde edilen bilgilerle Q-değerleri Q-Öğrenme güncelleme kuralı (Formül 2.1) ile güncellenir. Bu sayede, Dyna-Q gerçek deneyimlere ek olarak model üzerinde de öğrenme yapar ve böylece daha hızlı bir öğrenme ve adaptasyon olanağı sunar.



3. LİTERATÜR ÇALIŞMASI

3.1. Strateji Oyunları

RTS oyunlarının, stratejik karar verme süreçlerini simüle etme konusunda etkili bir araç olduğu literatürde yaygın biçimde kabul edilmektedir. Dachowicz vd. (2022), RTS tabanlı oyunların görev mühendisliği ve tasarıma uygulanmasını inceleyerek oyun ortamında yapay zeka yaklaşımlarının nasıl açıklanabilir hâle getirilebileceğini tartışmıştır. Benzer bir bakış açısıyla Frank (2022), strateji temelli savaş oyunlarında YZ unsurları olmadan bile oyun tasarımının ve karar mekanizmalarının nasıl geliştirilebileceğini ele almıştır.

Özellikle, CTF oyunları gibi senaryolar, saldırı ve savunma stratejilerinin test edilmesinde sıkça kullanılmaktadır. Balogh vd. (2022), CTF tabanlı bir güvenlik ortamında SQL enjeksiyon saldırıları için dinamik optimizasyon mekanizmaları geliştirerek strateji tabanlı oyunların siber güvenlik alanındaki potansiyelini ortaya koymuştur. Oyunlar ve simülasyonların stratejik analizlerde artan önemi, Goecks vd. (2023) çalışmasında da vurgulanmıştır; bu çalışmada, komuta ve kontrol uygulamaları için yapay zeka temelli yaklaşımların prototip aşamasından operasyonel düzeye nasıl taşınabileceği incelenmiştir.

Savaş sisinin strateji oyunlarında kritik bir unsur olduğu ve gerçeklik algısını güçlendirdiği, Obeia ve Wójcik (2023) tarafından ortaya konulmuştur. Dijital ortamlarda bu sis etkisinin simülasyonu, gerçek hayattaki belirsizliği modellese de optimizasyon ve karar verme süreçlerini karmaşıktırmaktadır. Askerî stratejilerle bağlantılı olarak, Cai ve Zhang (2023), belirsizliğin stratejik güven ve uluslararası ilişkiler açısından da kritik bir rol oynadığını savunmaktadır. Murray (2023) ise tarihsel bir bakış açısıyla, Clausewitz'in kart oyunlarına dair analogilerini

kullanarak askerî stratejide şans ve belirsizliğin nasıl değerlendirilmesi gerektiğini tartışmıştır.

Bununla birlikte, strateji oyunları yalnızca askerî veya siber güvenlik amaçları için değil, eğitim ve öğretim açısından da değerlidir. Örneğin, Cole (2022), CTF etkinliklerini bilgisayar güvenliği derslerinde kullanarak öğrencilerin öğrenme çıktılarını iyileştirmiştir. Açık kaynaklı CTF ortamlarının işlevsellikleri ve yapılandırmaları üzerine yapılan ampirik bir anket, bu tür ortamların eğitimciler ve araştırmacılar için hayati önemini vurgulamaktadır (Kucek & Leitner, 2020). Benzer şekilde, Minagar ve Sakzad (2023), IT Forensics alanında CTF-stil alıştırmaların otomatik problem üretimi üzerinde durarak, bu tarz oyunların eğitimsel potansiyelini ortaya koymuştur. Uzaktan eğitim araçlarından biri olarak sunulan savaş oyunlarının, öğrencilerin siyasî bilim gibi alanlarda katılımını artırdığı; sanal eğitimde belirsizlik ve karşılaşmaların beslendiği bir ortam sunduğu da literatürde belirtilmektedir (Harris & Freeman, 2023).

3.2. Pekiştirmeli Öğrenme

Literatürde, pekiştirmeli öğrenme yöntemlerinin strateji tabanlı oyunlarda son yıllarda giderek daha fazla kullanılmaya başlandığı görülmektedir. Silver vd. (2018), satranç, shogi ve Go gibi karmaşık oyunlarda pekiştirmeli öğrenme kullanarak insanüstü performans elde edilebileceğini göstermiştir. Bunu takiben, Silver vd. (2017), derin pekiştirmeli öğrenme (Deep RL) ve kendi kendine oynama (self-play) yöntemlerini birleştirerek AlphaGo Zero'nun Go oyununu insan bilgisine ihtiyaç duymadan ustalıkla oynamasını sağlamıştır. Bu çalışmalar, yapay zekâ algoritmalarının sadece tahmine dayalı değil, aynı zamanda stratejik karar alma becerilerini de geliştirilebilir hâle getirdiğini göstermektedir.

Benzer şekilde, saldırgan güvenlik eğitiminde de pekiştirmeli öğrenme yöntemleri etkin biçimde kullanılmıştır. Örneğin, Erdódi ve Zennaro (2022), web uygulaması güvenlik testlerinde saldırı senaryolarını modellemek için RL tabanlı bir yaklaşım önermiştir. CTF çerçevesini ele alan bir başka çalışmada, Chetwyn ve Erdódi (2022), Gym platformunu kullanarak SQL enjeksiyon saldırıları üzerine odaklanan dinamik eğitim ortamları tasarlamıştır. RTS oyunlarında ise büyük eylem alanları ve kısmi gözlemlenebilirlik gibi zorluklara rağmen derin pekiştirmeli öğrenmenin uygulanabilirliği sürmektedir Van Stolen vd. (2020): Adil vd. (2017).

Büyük ölçekli ortamlarda çoklu ajan yaklaşımını kullanan çalışmalara örnek olarak, Jaderberg vd. (2019a) tarafından geliştirilen Popülasyon Tabanlı Pekiştirmeli Öğrenme (Population-Based RL) modeli, 3D çok oyunculu ortamlarda insan seviyesinde performans elde etmeyi başarmıştır. Vinyals vd. (2019), Vinyal vd. StarCraft II gibi karmaşık RTS oyunlarında çoklu ajan RL yöntemlerini kullanarak rekabetçi sonuçlara ulaşmış; AlphaStar gibi modellerin grandmaster seviyesinde performans sergilemesini sağlamıştır (2019). Ancak, Adil vd. hâlâ RTS oyunlarında eksik gözlem, büyük durum-uzayları gibi temel zorlukların tam anlamıyla çözülemediği bildirilmektedir (2017).

Öğrenme stratejilerinin optimizasyonu ve farklı bağlamlardaki performans farklılıkları da literatürde dikkat çekmektedir. Zhong (2024), Q-Öğrenme ve SARSA algoritmalarını Cliff Walking problemi üzerinde karşılaştırarak, Q-Öğrenme'nin daha agresif bir yaklaşım sergilediğini, SARSA'nın ise daha temkinli sonuçlar ürettiğini göstermiştir. Zennaro ve Erdodi (2023), CTF senaryolarını kullanarak önceden bilgi sahibi olmanın (model-based) avantajlarını ve dezavantajlarını değerlendirmiştir. Benzer şekilde, Beason vd. (2024), denizcilik alanındaki CTF yarışmalarında iş birliğine dayalı otonominin düşmanca

ortamlardaki etkinliğini ortaya koymuştur. Kliem ve Dasgupta (2023), RTS oyunlarında ödül şekillendirmenin öğrenmeyi geliştirmedeki rolünü irdelerken, Smith ve Tanaka (2024) çoklu ajanlı rekabetçi oyunlarda sezgisel öğrenme stratejilerini ele almış, Doe ve Lee (2023) ise RL algoritmalarının dinamik strateji oyunlarındaki karşılaştırmalı performansını mercek altına almıştır.

Adversarial Networks gibi çatışmacı öğrenme yaklaşımları da strateji oyunlarında dikkate değerdir. Spencer vd. (2021), yapay zekâ tabanlı stratejik oyunların saldırgan koşullarda karar alma süreçlerini iyileştirme potansiyelini araştırmış, Jaderberg vd. (2019b) ise popülasyon tabanlı yöntemler kullanarak karmaşık 3D çok oyunculu ortamlarda yüksek performansa ulaşılabileceğini kanıtlamıştır. Son olarak, insan siber operatörlerini anlamak ve veri toplamak üzere CTF yarışmalarının eğitsel boyutu önem kazanmış; Savin vd. (2023) bu çerçevede veri etiketleme yöntemlerini geliştirmiştir. Farshi (2023), strateji oyunlarındaki şans faktörünü incelerken; bu durumun özellikle spor dallarında “şans kümelenmesi” yaratabileceğini öne sürmüştür.

Mevcut literatür incelendiğinde, yapay zekânın CTF ve benzeri strateji oyunlarındaki potansiyelinin oldukça yüksek olduğu; ancak belirsizlik ve şans faktörlerinin gerçekçi simülasyonlarda yeterince ele alınmadığı göze çarpmaktadır. Bu çalışma, savaş sisini ve şans faktörlerini stratejik simülasyonlara entegre ederek bu boşluğu doldurmayı amaçlamakta; böylece eksik bilgi ve stokastik olayların yapay zekâ temelli karar alma süreçlerini nasıl etkilediğine dair yeni içgörüler sunmaktadır.

4. YÖNTEM / SİSTEM ÖNERİSİ

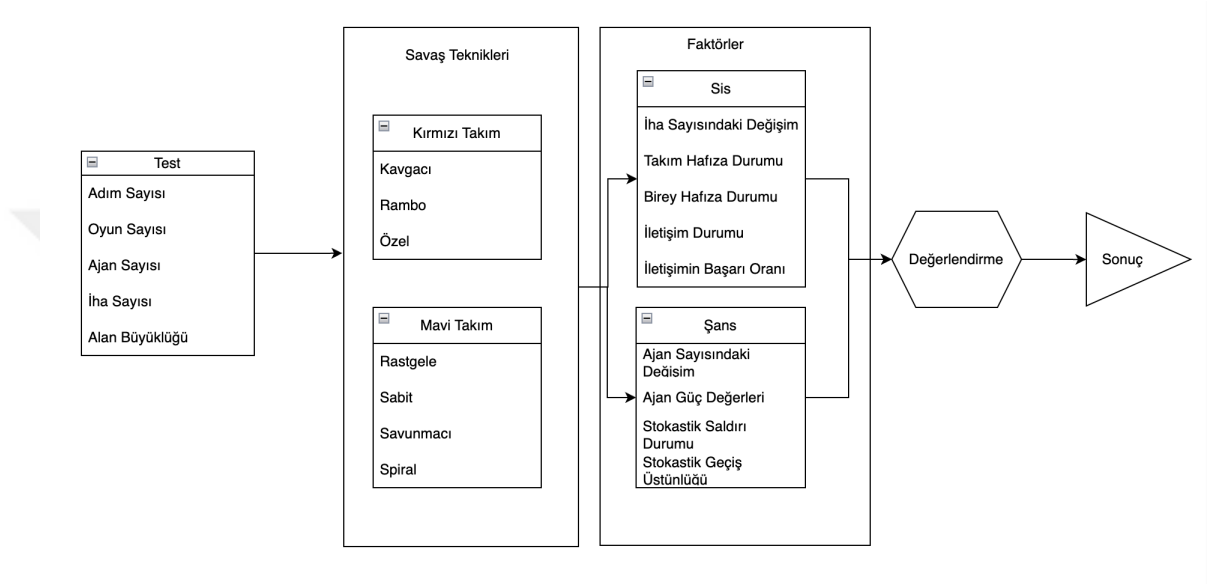
Bu çalışma, Gym-Cap ortamında sezgisel öğrenme yaklaşımları ile sis ve şans faktörlerini test etmeyi ve bunun yanı sıra, geliştirilmiş basit bir CTF oyunu üzerinde birkaç pekiştirmeli öğrenme modelini değerlendirmeyi amaçlamaktadır. Gym-Cap tabanlı sezgisel modeller, CTF oyun çerçevesinde sis ve şans faktörlerinin etkilerini test etmek üzere eğitilmiş; temel düzey CTF oyununda ise Q-Öğrenme ve diğer RL modelleri, otomatik parametre optimizasyonu ve sis ile şans faktörleri altında test edilmiştir. Sis ortamında şans faktörünü simüle etmenin amacı ve hedefi (1) oyunun gerçekçiliğini artırmak ve (2) oyunculara savaş koşullarında karşılaşabilecekleri belirsizliklere uyum sağlamayı öğretmektir. Sonuç olarak, bu çerçevede, teorik yapay zeka stratejileri ile pratik, gerçek dünya savaş senaryolarında uygulamaları arasında bir köprü kurmayı amaçlamaktadır. Sis ortamında şans faktörünü simüle etmek, oyuncuların stratejik düşünme yeteneklerini geliştirmeye yardımcı olma potansiyeline sahiptir. Örneğin, bir savaş senaryosunda, belirsiz hava koşulları nedeniyle bir hedefin doğruluk oranının azalması durumunda oyuncuların stratejilerini ve eylemlerini değiştirmeleri beklenmektedir. Bu şekilde, amaç, oyuncuların savaş koşullarında başarılı olmasını sağlayan gerçekçi bir savaş ortamını simüle etmektir. Önerilen yöntem/sistem, takımların öğrenilen davranışlara dayalı olarak stratejilerini uyarladığı ve zamanla performanslarını geliştirdiği stratejik bir savaş oyunu senaryosunu simüle etmeyi hedeflemektedir. Rastgele sayılar, şans faktörü olarak kullanılacaktır. Rastgele sayılar, şans faktörü kapsamında taraflara avantaj veya dezavantaj sağlayan olaylar yaratmak için kullanılmış; ayrıca temel düzey CTF oyunu için ajanların rastgele oyunu kazanma ya da ek puan kazanma gibi şans mekanizmaları eklenmiştir. Farshi (2023) tarafından incelenen strateji oyunlarındaki şansın performans ve strateji üzerindeki

etkileri önemli olup, spor dallarında 'şans kümelenmesi' olarak adlandırılmıştır. Askeri stratejide, şans ve belirsizliğin rolü tarihsel olarak kabul edilmiştir; Murray (2023), Clausewitz'in kart oyunlarına benzetmelerini kullanarak stratejistlerin şansını nasıl kullanabileceğini tartışmıştır. Savaş oyunlarında 'savaş sis' kavramını simüle etmek, savaş oyunlarının gerçekçiliğini artırmak için esastır. Obeia ve Wójcik (2023), dijital savaş oyunlarındaki farklı 'savaş sis' modellerini karşılaştırmış ve stratejik simülasyonlarda etkili sis modellerinin önemini vurgulamıştır.

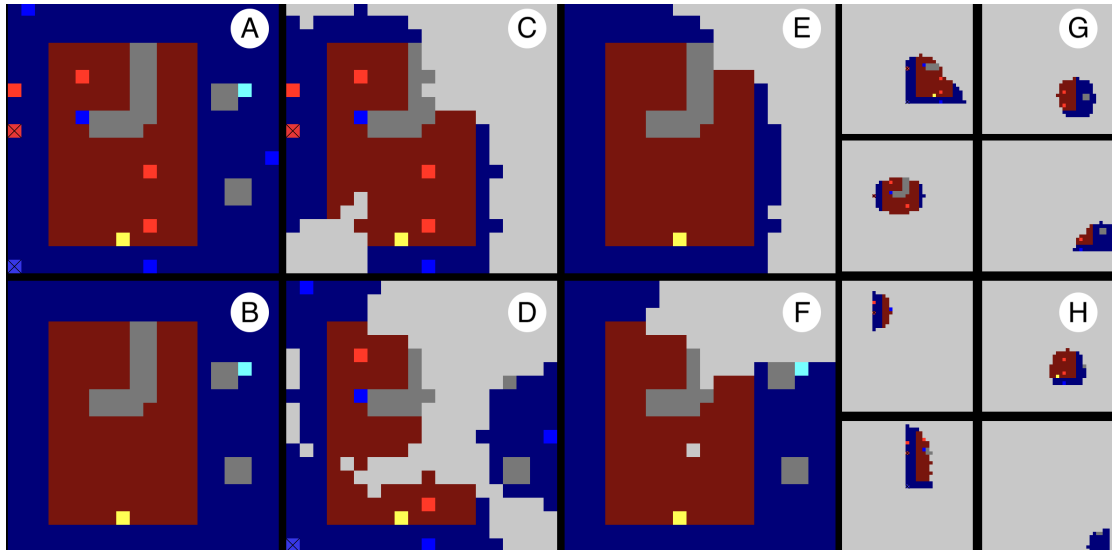
4.1. CTF Test Ortamı

Bu çalışmada, bağımsız çoklu ajan ortamlarında adaptasyonu araştırmak amacıyla Gym-Cap ortamı kullanılarak bir CTF Test Ortamı oluşturulmuştur. CTF Test Ortamı, ajanların işbirliği yapması, strateji geliştirmesi ve politikalarını başarıya ulaştırmak için uyum sağlaması gereken dinamik ve zorlu bir ortam sunmaktadır. Bu ortamda, mavi ve kırmızı olarak temsil edilen iki takım, belirli bir kapsama alanında saldırı ve savunma stratejileriyle mücadele etmektedir. Mavi takım savunma görevini üstlenirken, kırmızı takım saldırı rolünü üstlenmektedir. Bu düzenek, takımların eylemlerini koordine etmesi ve değişen koşullara uyum sağlaması gereken gerçek dünya senaryolarını simüle etmektedir. Ayrıca, sis ve şans faktörlerinin dahil edilmesi, ajanların sadece statik koşullar için optimize olmalarını değil, aynı zamanda gerçek dünya benzeri öngörülemezliklerle başa çıkmayı da öğrenmelerini sağlamaktadır. Kapsamlı bir değerlendirme sağlamak için, tüm denemelerde tutarlı parametreler kullanılarak bir dizi deney yapılacaktır. Özellikle, her deneme 1000 oyun içerecek ve her oyun 500 adımdan oluşacaktır. Kapsama alanı 20x20 kare olarak ayarlanacak, bu da ajanlar için dengeli ve zorlu bir oyun alanı sağlayacaktır. Şekil 4.1, test simülasyonunun akış diyagramını göstermektedir. Test

kısımındaki tüm testler boyunca sabit tutuldu. Savaş taktiklerinde kırmızı ve mavi ajanların uyguladıkları savaş stratejilerini göstermekte. Faktörler kısmında ise test edilen sis ve şans faktörleri belirtilmekte. Daha sonrasında çıktılar değerlendirilmeden geçirilerek sonuçlandırılmakta.



Şekil 4.1. CTF Simülasyon Çerçevesi Akış Diyagramı



Şekil 4.2. CTF Gym Tam Ekran. (a) Tüm oyuncular ve bayraklar. (b) Sadece bayraklar. (c)-(d) Takım görünümüleri. (e)-(f) Takım alanları. (g)-(h) Bireysel alanlar.

Şekil 4.2, CTF gym'in bir dizi piksel sanatı görüntüsünü göstermektedir. CTF gym, ajanların stratejilerini pratik yapabildikleri bir yerdir. Her takımın kendi alanı, bayrağı, oyuncuları ve gözlemcileri bulunmaktadır. Gri alanlar ortak paylaşılan alanları temsil etmektedir. Görüntüler, gym'in farklı bölümlerini temsil etmektedir. Şekil 4.2a, tüm oyuncuları ve bayrakları göstermektedir. Bu, CTF'nin başlangıç noktası olup, oyundaki tüm oyuncuları ve buldukları bayrakları göstermektedir. Şekil 4.2c ve 4.2d, takım oyuncularını ve görebildikleri alanları göstermektedir. Bu görüntüler, aynı takımda bulunan oyuncuları ve takımın görebildiği harita alanlarını göstermektedir. Şekil 4.2e ve 4.2f, takımın keşfettiği harita alanlarını göstermektedir. Bu görüntüler, takım tarafından keşfedilen harita alanlarını ve bu alanlarda bulunan bayrakları göstermektedir. Şekil 4.2g ve 4.2h, bireylerin keşfettiği harita alanlarını göstermektedir. Bu, bireysel oyuncular tarafından keşfedilen harita alanlarını ve bu alanlarda bulunan bayrakları göstermektedir. Şekil 4.2b, oyuncular olmadan sadece bayrakları göstermektedir. Bu görüntü, bulunan bayrakları belirlemek ve sonraki hareket planını yapmak için faydalıdır.

4.2. Öğrenme Metodolojisi: Yapay Zeka Ajanları

Bu çalışmada kullanılan öğrenme metodolojisi, Gym-Cap platformu içinde sezgisel öğrenme algoritmalarının gücünden yararlanarak ajanların uyarlanabilir, deneyime dayalı kararlar almasını sağlamaktadır. Sezgisel öğrenme, diğer yapay zeka öğrenme tekniklerinden, örneğin derin pekiştirmeli öğrenme veya sinir ağları gibi, birkaç önemli şekilde ayrılır. Sezgisel öğrenme, insan sezgisini taklit eden basitleştirilmiş akıl yürütme modellerine odaklanır, bu da geniş veri setleri veya yüksek hesaplama kaynaklarına ihtiyaç duymadan daha hızlı karar almayı mümkün kılar. Bu hız ve basitlik, hızlı uyum sağlamanın kritik olduğu gerçek zamanlı strateji oyunlarında

avantajlıdır. İkinci olarak, sezgisel öğrenme, alan bilgisine dayalı net, kural tabanlı stratejilere dayandığından daha fazla açıklanabilirlik sunar. Bu şeffaflık, yapay zekanın karar alma süreçlerinin anlaşılmasını ve analiz edilmesini kolaylaştırır, bu da stratejilerin iyileştirilmesi ve oyun tasarımının geliştirilmesi için faydalıdır. Ayrıca, sezgisel öğrenme, sınırlı veya eksik veriye sahip ortamlara iyi uyum sağlar, çünkü uzun eğitim sürelerine ihtiyaç duymadan etkili bir şekilde çalışabilir. Dinamik ve belirsiz ortamlarda – örneğin, savaş sis etkileri ve şans faktörleri ile – sezgisel öğrenme, ajanların sezgisel kuralları ve geçmiş deneyimleri uygulayarak yeni durumlara hızla uyum sağlamasına olanak tanır. Bu, derin öğrenme gibi yöntemlerin yeniden eğitim ve yeni koşullara uyum sağlama konusunda önemli zaman gerektirebileceği durumların aksine bir avantajdır. Genel olarak, sezgisel öğrenme, belirsizlik ve sınırlı bilgiyle başa çıkabilen stratejik oyunların karmaşıklıklarını yönetebilecek yapay zeka ajanları geliştirmek için pratik ve etkili bir yaklaşımdır.

Bu çalışmada Gym-Cap olarak adlandırılan bir çerçeve kullanılmıştır. Bu çerçeve, OpenAI Gym kütüphanesine eklenmiş bir genişletme paketidir ve CTF oyunlarını simüle etmek için tasarlanmıştır. Bu çerçevede, model eğitimi sırasında başarının değerlendirilmesi için çeşitli metrikler kullanılmıştır. Bu metrikler, CTF oyununda elde edilen puanlar, oyunda hayatta kalma süresi ve düşman ajanlardan kaçınma oranı gibi performans kriterlerini içerir. Ayrıca, model eğitim sürecinde ödül eğrileri, hata oranları ve algoritmanın yakınsama davranışı kaydedilmiştir. Özellikle, model öğrenme sürecinde pekiştirmeli öğrenme yöntemleri kullanılmıştır. Bu yaklaşım, ajanların çevre ile etkileşime girip aldıkları eylemler üzerine ödül veya ceza almasını sağlar. Yaygın olarak kullanılan pekiştirmeli öğrenme teknikleri arasında Q-öğrenme, derin Q-ağları (DQN), politika gradyanı yöntemleri ve aktör-eleştirmen yöntemleri bulunmaktadır. Bu teknikler, ajanların çevreye etkili

bir şekilde uyum sağlamalarını sağlamak için ödül maksimizasyonu stratejileri geliştirmelerini teşvik eder. Ayrıca, rekabetçi eğitim yöntemleri, ajanların birbirlerinden öğrenmesini sağlamak için uygulanmıştır.

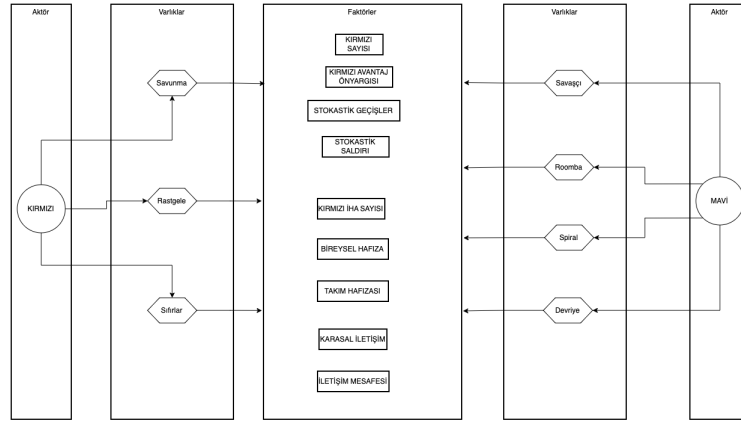
Öncelikle Gym-Cap tabanlı bir test ortamında, sis ve şans faktörlerinin yapay zekâ ajanlarının performansına etkileri incelenmiştir. Bu ortam, çok ajanlı dinamik senaryolar sunarak stratejik karar alma süreçlerinin çeşitli yönlerini değerlendirme imkânı tanımıştır. Ancak Gym-Cap'ın belirli kısıtları (bakım eksikliği, saldırgan stratejileri uygulamadaki zorluklar) nedeniyle, ek olarak daha temel düzeyde ve kontrol edilebilir bir CTF ortamı oluşturulmuştur. Bu ikinci ortamda, Q-Öğrenme, SARSA, Çift Q-Öğrenme gibi farklı RL algoritmaları karşılaştırılmış, parametre optimizasyonu, belirsiz bilgi koşulları ve rastgele ödüllerin etkisi daha yalın bir biçimde analiz edilmiştir.

4.3. Test Metodolojisi

Test metodolojisi, Gym-Cap CTF test ortamında yapay zeka ajanlarının performansını değerlendirmek için kapsamlı bir yaklaşım içermektedir. Sistemik deneyler aracılığıyla, ajan davranışlarının ve dinamik savaş oyunları senaryolarındaki uyum yeteneklerinin çeşitli yönleri incelenmektedir. Bu, bayrak yakalama ve savunma stratejilerinde hiyerarşik stratejilerin etkinliğinin değerlendirilmesini, ajanların belirsiz ve dinamik ortamlara (örneğin, simüle edilmiş savaş sis koşulları) uyum sağlama yeteneklerinin analiz edilmesini ve farklı soyutlama seviyelerindeki karar verme yeteneklerinin değerlendirilmesini içermektedir. Ayrıca, ajanlar arasındaki etkileşim dinamikleri ve takım içi koordinasyonları, iletişim ve işbirliği stratejileri gibi faktörler de incelenmektedir. Şans faktörleri ve rastgele olayların oyun sonuçları üzerindeki etkisi de odak

noktasıdır, bu da ajanların öngörülemez koşullara uyum sağlama yeteneklerini ölçmeyi amaçlamaktadır. Zaman bazlı kısıtlamalar ve savaş sis etkilerine yapılan ayarlamalar, ajanların zaman baskısı ve eksik bilgi senaryolarında karar verme etkinliğini test etmeye olanak tanımaktadır. Gym-Cap'in esnek performans metrikleri ve değerlendirme kriterleri kullanılarak, tekrarlamalı testler ve iyileştirmeler yapılmakta, böylece yapay zeka ajanlarının stratejik savaş oyunları senaryolarındaki yetenekleri sürekli olarak geliştirilmiştir. Bu çalışmada, her bir parametre değiştirildiğinde diğerleri orijinal durumuna sıfırlanmış, böylece her bir parametrenin bağımsız olarak etkisi gözlemlenebilmiştir.

Çerçeve, CTF simülasyonunda stratejik karar alma süreçlerini etkileyen aktörlerin, varlıkların ve temel faktörlerin rolü incelenmektedir. Şekil 4.3, bu çerçevede çevresel değişkenler ile ajan davranışları arasındaki etkileşimi göstermektedir. Örneğin, 'KIRMIZI AJAN' ve 'KIRMIZI AVANTAJ ÖNYARGISI' arasındaki etkileşim, şans faktörlerinde küçük ayarlamaların her iki takımın da stratejilerini nasıl önemli ölçüde değiştirebileceğini göstermektedir.



Şekil 4.3. Çerçeve Aktörler, Varlıklar ve Faktörler

Algoritma 1, CTF senaryolarında mavi ve kırmızı takımların kullanacağı strateji ve

taktikleri belirleyerek bu alandaki derinlemesine anlayışımızı artırmaktadır. Örneğin, farklı politika kombinasyonları kullanarak yapılan maç simülasyonları sayesinde bu algoritma, belirsiz koşullar altında her iki takım için de en uygun savunma ve saldırı stratejilerini belirlemeye yardımcı olur. Çeşitli politika kombinasyonlarını kullanarak yapılan kapsamlı simülasyonlar sayesinde, bu algoritma saldırı ve savunma stratejileri arasındaki etkileşimi ve bunların genel performans üzerindeki etkilerini ortaya koyar. Böylece, dinamik ve rekabetçi ortamlarda bilinçli kararlar almak için içgörüler sağlar.



Algoritma 1: GYM-Cap ile Strateji Simülasyonu Algoritması

Veri: Red_list, Blue_list, Environment object, Policy object, output_file

Sonuç: Her kombinasyon için Ortalama Süre ve Ortalama Puanı Yazdır

$bl \leftarrow 0;$

$rl \leftarrow 0;$

$observation \leftarrow env.reset(configuration_file);$

$num_match \leftarrow 1000;$

$render \leftarrow False;$

$action \leftarrow env.action_space.sample();$

$rscore \leftarrow [];$

$start_time \leftarrow time.time();$

while bl in Blue_list **do**

while rl in Red_list **do**

$done \leftarrow False;$

$rewards \leftarrow [];$

Döngü n in range(num_match) **yap**

while not $done$ **do**

$observation, reward, done, info \leftarrow env.step(action);$

$rewards.append(reward);$

$env.reset();$

$rscore.append(sum(rewards));$

$duration \leftarrow time.time() - start_time;$

$average_time \leftarrow duration/num_match;$

$average_score \leftarrow sum(rscore) / num_match;$

return("kırmızı: " + rl + " mavi: " + bl + " Ortalama Süre:

4.3.1. Gelişmiş test yöntemi

Bu çalışmada, Gym-Cap tabanlı testlerin yanı sıra, ek olarak daha temel düzeyde bir CTF oyun ortamında da yapay zeka ajanlarının performansları değerlendirilmiştir. Bu temel düzey CTF ortamı, Gym-Cap gibi çoklu ajan ve karmaşık strateji unsurlarına sahip olmakla birlikte, daha küçük bir ızgara (10x10) ve daha sınırlı sayıda kural çerçevesinde tasarlanmıştır. Bu sayede, ajanların temel öğrenme dinamiklerini ve uyum yeteneklerini daha yalın bir ortamda test etmek mümkün olmuştur. Burada, saldırgan ve savunmacı olmak üzere iki temel rol tanımlanmış; saldırganın amacı bayrağı bulmak, savunmacının amacı ise saldırganı yakalamak şeklinde belirlenmiştir.

Bu yeni test altyapısında, Q-Öğrenme, SARSA, Çift Q-Öğrenme, Beklenen SARSA, Monte Carlo ve Dyna-Q gibi farklı pekiştirmeli öğrenme algoritmaları aynı temelde karşılaştırılmıştır. Algoritma 2’de örnek bir eğitim ve test fonksiyonu yer almaktadır. Bu fonksiyon, her eğitim bölümünde belirli sayıda adım atmakta, tüm modellerin durum-eylem çiftlerini güncellemekte, CPU ve RAM kullanımı gibi sistem kaynak tüketimlerini izlemekte ve belirli aralıklarla (örneğin, her 1000 bölümde bir) testler yaparak ajanların ortalama ödül, adım başına performans ve zaman kullanımı gibi metriklerini kaydetmektedir. Ayrıca, “sis” faktörü etkinleştirilerek, ajanların kısıtlı bilgi koşullarında nasıl performans gösterdiği; “şans” faktörü devreye alınarak ise rastgele ek puanlarla ajan stratejilerinin ne yönde etkilendiği incelenmiştir.

Böylece, hem Gym-Cap test ortamında yapılan kontrollü deneyler hem de daha temel düzey bir CTF uygulaması ile gerçekleştirilen kapsamlı testler sayesinde, ajanların farklı koşullar altındaki strateji uyarılma yetenekleri daha geniş bir perspektiften değerlendirilmiştir. Bu yaklaşım, hem karmaşık hem de basit senaryolarda elde edilen bulguların bütüncül bir biçimde yorumlanmasına imkân

tanınmaktadır.

Algoritma 2: Temel Düzey CTF Ortamı için Eğitim ve Test Fonksiyonu

Veri: Aksiyon seti, ajan modelleri, bölüm sayısı, test aralığı, adım sınırı

Sonuç: Her model için eğitim süresince toplanan ortalama ödüller,

CPU/RAM kullanımı, adım sayısı, vb.

Döngü $episode \leftarrow 1$ **to** $total_episodes$ **yap**

Ortamı sıfırla ve ajanların durumunu al;

while *not done* **and** $step_count < max_steps$ **do**

Ajanların mevcut bilgisine (sis) ve rastgelelik koşullarına (şans) göre
eylemleri seç;

Eylemleri ortamda uygula, ödülleri ve yeni durumları hesapla;

Ajanların öğrenme fonksiyonlarını çağırarak Q-değerlerini güncelle;

CPU ve RAM kullanımını ölç, kaydet;

Eğer $episode \% test_interval == 0$ ise: Eğitimli ajanları rastgele
ajanlarla test et, ortalama ödülleri hesapla;

Performans metriklerini kaydet (zaman, adım, CPU, RAM);

4.3.2. Sis ve şans faktörleri

Test metodolojisinde rastgelelik ve bilgi kısıtlamalarını içeren şans ve sis faktörleri, ajanların belirsiz ve dinamik koşullardaki strateji uyarlama yeteneklerini değerlendirmek için özel olarak tasarlanmıştır. Şans faktörü, oyun ortamına rastgelelik eklemek amacıyla iki mekanizma ile gerçekleştirilmiştir. Ajanların her bir adımda 1/100 ihtimalle doğrudan oyunu kazanabilmesi sağlanmıştır. Bu mekanizma, düşük olasılıkla gerçekleşen ancak yüksek etkili rastgele olayların oyun üzerindeki etkisini modellemek amacı taşımaktadır. İkinci olarak, her bir adımda 1/10 ihtimalle bir ajanın +0.1 puan ekstra kazanabilmesi sağlanmıştır. Bu durum

ise daha sık gerçekleşen ancak daha düşük etkili rastgele ödüllerin, ajan davranışlarını ve stratejik kararlarını nasıl etkilediğini inceleme fırsatı sunmuştur. Bu iki yapı, şans faktörünün stratejik bir oyun ortamında ödüllere, kazançlara ve ajanların öğrenme süreçlerine nasıl etki ettiğini değerlendirmeyi mümkün kılmıştır.

Sis faktörü, ajanların çevresel bilgi seviyelerini sınırlayarak eksik bilgi koşulları altında strateji geliştirme yeteneklerini analiz etmek için uygulanmıştır. Üç farklı bilgi seviyesi tanımlanarak ajanların farklı senaryolarda nasıl performans gösterdiği incelenmiştir. İlk senaryoda, ajanlar yalnızca kendi konumlarını görebilmekte, diğer ajanların ve bayrağın konumları hakkında herhangi bir bilgiye sahip olmamaktadır. Bu durum, bilgi eksikliği koşullarını simüle ederek ajanların sadece kendi durumlarına dayalı kararlar almasını zorunlu kılmaktadır. İkinci senaryoda, ajanlara eksiksiz bilgi sağlanmış ve hem kendi konumlarını hem de diğer tüm varlıkların (bayrak ve diğer ajanlar) konumlarını görebilmeleri sağlanmıştır. Bu, tam bilgiye dayalı stratejik kararların alınabileceği ideal bir senaryoyu temsil etmektedir. Üçüncü senaryoda ise ajanlar yalnızca kendi konumlarını ve hedeflerinin (saldırgan için bayrak, savunmacı için saldırgan) konumlarını görebilmektedir. Bu bilgi seviyesi, eksik bilgi ve tam bilgi arasında bir denge kurarak, daha gerçekçi bir stratejik ortam sunmaktadır.

Şans ve sis faktörlerinin bu şekilde modellenmesi, yapay zekâ algoritmalarının hem öngörülemeyen koşullara adapte olma hem de eksik bilgi altında strateji geliştirme kapasitelerini değerlendirme açısından önemli katkılar sunmuştur. Bu faktörler, stratejik karar alma süreçlerini ve öğrenme mekanizmalarını daha kapsamlı bir şekilde analiz etmeye olanak sağlamış, özellikle belirsiz koşullarda yapay zeka performansının nasıl değiştiğini anlamaya yönelik değerli içgörüler ortaya koymuştur.

4.4. Temel Düzey CTF

Algoritma 3, 10x10 boyutunda bir ızgara üzerinde oynanan basit bir CTF oyununu modellemektedir. Oyun, saldırgan ve savunmacı olmak üzere iki oyuncudan oluşur. Saldırganın amacı bayrağı bulmak, savunmacının amacı ise saldırganı yakalamaktır. Bayrak ve ajanların başlangıç konumları rastgele belirlenir. Her iki ajan da beş hareket seçeneğine sahiptir: yukarı, aşağı, sağ, sol ve bekleme. Algoritmada, her adımda ajanlar mevcut durumlarına göre bir eylem seçer ve konumları bu eylemler doğrultusunda güncellenir. Ardından, belirlenen ödül mekanizmasına göre puanlar güncellenir.

Oyun, iki temel durumda sona erer: saldırgan bayrağı bulduğunda veya savunmacı saldırganı yakaladığında. Saldırgan bayrağı bulduğunda, saldırgan +20 puan kazanırken savunmacı 5 puan kaybeder. Tersine, savunmacı saldırganı yakaladığında, savunmacı +20 puan kazanırken saldırgan 5 puan kaybeder. Bunun yanı sıra, ajanların hedeflerine olan mesafelerindeki değişiklikler de ödülleri etkiler. Örneğin, saldırgan bayrağa yaklaştığında +3 puan kazanır, uzaklaştığında ise 2 puan kaybeder. Benzer şekilde, savunmacı saldırganı yaklaştığında +3.5 puan kazanırken uzaklaştığında puan kaybeder. Bu ödül sistemi, ajanların yalnızca hedefe odaklanmak yerine çevresel dinamikleri de dikkate almasını sağlar.

Algoritmada 3, isteğe bağlı olarak bir şans faktörü de bulunmaktadır. Şans faktörü etkinleştirildiğinde, ajanlara her adımda rastgele 0 ile 4 arasında ek puan verilir. Bu, oyun mekaniğine rastgelelik katarak ajanların beklenmedik durumlara nasıl uyum sağladığını test eder. Ayrıca, sis faktörü devreye alındığında, ajanlar yalnızca kendi konumlarını gözlemleyebilir ve bayrak ya da rakiplerinin konumuna dair bilgi sahibi olamaz. Bu, ajanların eksik bilgiyle karar verme ve strateji geliştirme yeteneklerini değerlendirmek için ek bir zorluk oluşturur.

Genel olarak, algoritma 3, hem saldırgan hem de savunmacı için stratejik bir mücadele alanı sunar. Ajanların ödülleri ve konumları üzerinden oyun içi dinamikler belirlenir ve belirli bir maksimum adım sayısına ulaşıldığında veya oyun bitirme koşulları sağlandığında sonuçlar döndürülür. Bu yapı, stratejik yapay zeka algoritmalarını değerlendirmek için etkili ve anlaşılır bir test ortamı sağlamaktadır.

Algoritma 3: Temel Düzey CTF Oyunu

Veri: Izgara boyutu $N \times N$, Maksimum adım sayısı T_{\max}

Sonuç: Saldırgan ve Savunmacı ödülleri

Başlat: $A_{\text{pos}} \leftarrow$ saldırgan başlangıç konumu, $D_{\text{pos}} \leftarrow$ savunmacı başlangıç konumu, $F_{\text{pos}} \leftarrow$ bayrak konumu, $t \leftarrow 0$;

while $t < T_{\max}$ *ve oyun bitmedi* **do**

 Her ajan bir eylem seçer: Yukarı, Aşağı, Sağ, Sol veya Bekle;

 Konumlar eyleme göre güncellenir;

if *Saldırgan bayrağı bulursa* **then**

$R_A \leftarrow +20$, $R_D \leftarrow -5$, **Oyunu bitir**;

else if *Savunmacı saldırganı yakalarsa* **then**

$R_D \leftarrow +20$, $R_A \leftarrow -5$, **Oyunu bitir**;

else

 Mesafe değişimlerine göre ödülleri güncelle;

if *Şans faktörü etkinse* **then**

 Rastgele puan (0-4) ekle;

return *Ajan ödülleri ve oyun sonucu*;

5. UYGULAMA

Bu çalışma, stratejik savaş oyunlarındaki gerçekliği artırmak amacıyla iki aşamalı bir uygulama süreci içermektedir. İlk aşamada, RAIDE projesi kapsamında geliştirilen ve Gym-Cap kütüphanesi üzerine inşa edilen özelleştirilmiş bir CTF simülasyon paketi kullanılmıştır. İkinci aşamada ise, Gym-Cap dışında daha temel düzeyde ve özelleştirilmiş bir CTF oyunu geliştirilmiştir. Bu iki uygulama, yapay zeka ajanlarının farklı ortam ve koşullar altında performanslarını değerlendirmek ve optimize etmek amacıyla tasarlanmıştır.

5.1. Gym-Cap Uygulaması

Bu kavram kanıtı uygulaması, RAIDE projesi altında geliştirilen özelleştirilmiş bir CTF simülasyon paketini kullanarak gerçekleştirilmiştir (https://raide-project.github.io/ctf_research/). Bu paket, orijinal Gym-Cap kütüphanesi üzerine inşa edilmiştir ve OpenAI Gym platformuna bağımlılığı kaldırılmış, Gymnasium paketi entegre edilmiştir. Bu iyileştirme, yapay zeka modellerinin eğitim ve test edilmesi için daha fazla esneklik ve sağlamlık sunmakta, modern pekiştirmeli öğrenme çerçeveleri ile uyumluluğu sağlamaktadır. Gymnasium'a geçiş, özellikle strateji oyunlarının gerçekçi unsurlarının entegre edilmesi ve karmaşık simülasyon ortamlarının daha verimli bir şekilde yönetilmesi açısından kritik öneme sahiptir.

Ana hedef, sezgisel öğrenme algoritmalarının sis ve şans faktörlerini kullanarak strateji oyunlarının gerçekliğini nasıl artırabileceğini değerlendirmektir. RAIDE paketi, bu faktörlerin sistematik olarak uygulanıp test edilebileceği uyarlanabilir bir ortam sağlamıştır. Önerilen sistemin etkinliğini daha iyi değerlendirmek için

Gym-Cap ortamı kullanılarak kapsamlı simülasyonlar yapılmış, performans çeşitli metrikler aracılığıyla ölçülmüş ve sonuçlar mevcut yaklaşımlarla karşılaştırılmıştır. yapay zekanın adaptasyon kabiliyetini test etmek için PPO MlpPolicy algoritması 300.000 adım boyunca eğitilmiştir. Şeffaflık ve tekrarlanabilirlik sağlamak adına, çalışma için geliştirilen tüm kodlar proje deposunda halka açık olarak mevcuttur https://github.com/yonur65/ctf_public.

Deneysel kurulum, kırmızı ve mavi takımların 500 adım içinde 1000 oyun oynadığı 20x20 karelik bir ızgarada gerçekleştirilen simülasyonları içermiştir. Kırmızı takım genellikle saldırgan stratejiler kullanırken, mavi takım savunma manevralarına odaklanmıştır. Bu etkileşimler, ajanların dinamik ve öngörülemez koşullara uyum sağlama yeteneklerini değerlendirmek için tasarlanmıştır. Bu unsurlar, çevresel değişkenlikleri emüle etmek için dikkatlice entegre edilmiştir. Sis faktörü, sınırlı iletişim aralıkları, azalmış görünürlük ve bireysel ve takım belleklerine olan bağımlılık gibi kısıtlamalar aracılığıyla modellenmiştir. Bu kısıtlamalar, ajanların eksik bilgiyle çalışmasını zorunlu kılarak dinamik ortamlara uyum sağlama yeteneklerini test etmiştir. Şans faktörleri ise, rastgele ajan kayıpları, önyargılı takviyeler ve belirli eylemler için olasılıksal sonuçlar gibi stokastik unsurlar aracılığıyla entegre edilmiştir. Örneğin, ajan sayısı 4'ten 6'ya çıkarılmış ve stratejik karar alma süreçlerinde gerçekçi rastgelelikleri simüle etmek için stokastik geçişler ve saldırılar etkinleştirilmiştir.

Tablo 5.1, simülasyon ortamının karmaşıklığını ve gerçekçiliğini artırmak için yapılan parametre ayarlamalarını özetlemektedir. Bu ayarlamalar, iletişim mesafesinin 0'dan 2'ye artırılması, takım ve bireysel bellek yeteneklerinin etkinleştirilmesi ve yer iletişiminin eklenmesini içermektedir. Ayrıca, önyargı değeri (RED_ADV_BIAS) 0'dan 2'ye çıkarılmış, bu da bir takım için daha güçlü bir

avantajı temsil etmekte ve yapay zeka modellerinin adaptasyon yeteneklerini daha fazla test etmektedir. Bu deęişiklikler, ajanların belirsizlik ve rastgelelik içeren senaryolarda stratejilerini sürekli olarak yeniden deęerlendirmelerini ve uyarlamalarını gerektiren zorlu bir ortam sağlamıştır.

Çizelge 5.1. Gerçekçilięi ve Karmaşıklık İartırmak İçin Simülasyon Parametre Ayarlamaları

Tür	Deęişken	Eski	Yeni
Şans	Ajan sayısını artırma (NUM_RED)	4	6
Şans	Önyargı deęerini artırma (RED_ADV_BIAS)	0	2
Şans	Stokastik geçişleri artırma (STOCH_TRANSITIONS)	Doęru	Yanlış
Şans	Stokastik saldırıları artırma (STOCH_ATTACK)	Yanlış	Doęru
Sis	Hava ajan sayısını artırma (NUM_RED_UAV)	1	3
Sis	Takım belleęini etkinleştirmeye (INDIV_MEMORY)	Yanlış	Doęru
Sis	Bireysel belleęi etkinleştirmeye (TEAM_MEMORY)	Yanlış	Doęru
Sis	Yer iletişimini etkinleştirmeye (COM_GROUND)	Yanlış	Doęru
Sis	İletişim mesafesini artırma (COM_DISTANCE)	0	2

5.2. Temel Düzey CTF Oyunu

Bu çalışmada, Gym-Cap dışında daha temel düzeyde ve özelleştirilmiş bir CTF oyunu geliştirilmiş ve altı farklı yapay zeka modeli karşılaştırılmıştır. Bu temel düzey CTF oyunu, iki oyunculu bir yapıdadır ve saldırgan ve savunmacı olmak üzere iki ajandan oluşmaktadır. Oyunda, saldırganın amacı bayrağı bulmak, savunmacının amacı ise saldırganı yakalamaktır. Oyun, 10x10 boyutlarında bir ızgara üzerinde oynanmaktadır. Saldırgan, ilk satırda rastgele bir konumda başlarken, savunmacı sol sütunda rastgele bir konumda başlamaktadır. Bayrak ise ızgara üzerinde rastgele bir yerde bulunmaktadır.

Ajanlar, yukarı, aşağı, sağ, sol ve yerinde durma olmak üzere beş hareket yeteneğine sahiptir. Saldırgan, bayrağı bulduğunda +20 puan kazanırken, savunmacı 5 puan kaybetmektedir. Benzer şekilde, savunmacı saldırganı yakaladığında, savunmacı +20 puan kazanırken, saldırgan 5 puan kaybetmektedir. Eğer saldırgan, bayrağa yaklaşırsa +3 puan kazanmakta, uzaklaşırsa 2 puan kaybetmektedir. Mesafe değişmezse ve savunmacı saldırgana yaklaşırsa, saldırgan 1 puan kaybetmektedir. Savunmacı için de benzer şekilde, saldırgana yaklaştığında +3.5 puan kazanmakta, uzaklaştığında puan kaybetmektedir.

Bu temel düzey CTF oyununda, Q-Öğrenme, SARSA, Çift Q-Öğrenme, Beklenen SARSA, Monte Carlo ve Dyna-Q olmak üzere altı farklı yapay zeka modeli karşılaştırılmıştır. Modellerin karşılaştırılmasında, kullandıkları CPU ve RAM miktarı, atılan adım sayısı ve harcanan süre gibi performans metrikleri değerlendirilmiştir. 5000 oyunluk bir test yapılmış, her 500 adımda bir eğitilmiş saldırgan ve savunmacı ajanlar, eğitilmemiş modellerle 500 oyunluk bir testte karşılaştırılmıştır. Her oyunda maksimum 200 adımlık oyunlara izin verilmiştir.

Yapılan deneyler sonucunda, Q-Öğrenme algoritmasının en iyi performansı gösterdiği tespit edilmiştir. Bu modelin parametrelerini otomatik olarak ayarlayan bir kod eklenmiştir. Oyunda şans faktörünü simüle etmek için, ajanlara rastgele ek puanlar veren bir yapı eklenmiştir; verilecek puan 0 ile 4 puan arasında değişmektedir. Sis faktörünü test etmek için ise, ajan eğitimlerinde sadece kendi konumlarına dair bilgi verilmiş, diğer ajanların veya bayrağın konumu hakkında bilgi verilmemiştir.

5.2.1. Deneysel kurulum ve sonuçlar

Deneyler kapsamında, her model için 5000 oyun oynanmış ve her 500 adımda bir eğitilmiş ajanlar, eğitimsiz modellerle 500 oyunluk testlere tabi tutulmuştur. Her oyunda maksimum 200 adımlık oyunlara izin verilmiştir. Yapılan deneyler sonucunda, Q-Öğrenme algoritmasının en iyi performansı gösterdiği belirlenmiştir. Bu model için parametrelerin otomatik olarak ayarlanmasını sağlayan bir kod geliştirilmiştir. Tüm testler, MSKU Üniversitesine tahsis edilmiş bir bilgisayar üzerinde gerçekleştirilmiştir. Söz konusu sistemde 32 GB RAM, Intel(R) Xeon(R) CPU E5-2640 v4 @ 2.40GHz (40 çekirdek), 8 GB bellekli NVIDIA Quadro M4000 GPU ve Linux stargate 5.15.0-87-generic #97 20.04.1-Ubuntu işletim sistemi bulunmaktadır.

5.2.2. Şans ve sis faktörlerinin entegrasyonu

Oyunda şans faktörünü simüle etmek amacıyla, ajanlara rastgele ek puanlar veren bir mekanizma eklenmiştir. Bu puanlar 0 ile 4 arasında rastgele değerler almaktadır. Sis faktörünü test etmek için ise, ajanların eğitim sürecinde sadece kendi konumlarına dair bilgi verilmiş, diğer ajanların veya bayrağın konumu

hakkında herhangi bir bilgi sunulmamıştır. Bu sayede, ajanların eksik bilgi altında karar verme ve strateji geliştirme yetenekleri değerlendirilmiştir.



6. SONUÇLAR

Yapılan deneyler, sezgisel öğrenmenin karmaşık senaryolara uyum sağlama potansiyelini göstermiş ve stratejik simülasyonlarda sis ve şans faktörlerinin entegrasyonunun önemini pekiştirmiştir. Bu gelişmeler, oyunların gerçekçiliğini artırmakla kalmayıp, eğitim simülasyonları ve eğitici araçlar için de daha geniş uygulama alanlarının açılmasını sağlamaktadır.

Bu çalışmanın sonuçları, Tablo 6.3, 6.4 ve 6.5'da detaylandırılmaktadır ve sis ve şans faktörlerinin CTF simülasyonlarında yapay zeka modellerinin performansını nasıl etkilediğini göstermektedir. Parametreler ve kırmızı takımın "Ortalama Puan" üzerindeki etkileri, oyun sonuçlarını nasıl şekillendirdiği konusunda değerli içgörüler sunmaktadır ve bu bulgular, stratejik karar alma süreçlerinde stokastik ve çevresel dinamiklerin önemini vurgulamaktadır. Her bir parametreyi bağımsız olarak test ettiğimiz için, sis ve şans faktörlerinin oyun sonuçları üzerindeki bireysel etkilerini net bir şekilde izole edebilmiş ve gözlemleyebilmiş olduk.

Tablo 6.3'te görüldüğü üzere, NUM_RED parametresi, kırmızı takımın ortalama puanında -0.67'den -0.41'e önemli bir iyileşme göstermiştir. Bu, daha fazla ajanın eklenmesinin net bir avantaj sağladığını ve takımın belirsizlikle karakterize edilen senaryolarda uyum yeteneğini artırdığını göstermektedir. Benzer şekilde, RED_ADV_BIAS parametresi -0.87'den -0.41'e yükselmiş olup, şans faktörünün kırmızı takım için daha faydalı bir etki yarattığını göstermektedir. Ancak, STOCH_TRANSITIONS'ın etkisi minimal olup, -0.42'den -0.41'e sadece hafif bir iyileşme göstermiştir. Bu, şans olaylarının sonuçları etkilediğini, ancak esas olarak takım tepkilerini istikrara kavuşturarak performansı önemli ölçüde değiştirmediklerini göstermektedir. STOCH_ATTACK parametresi ise, ortalama puanın -0.41'de sabit

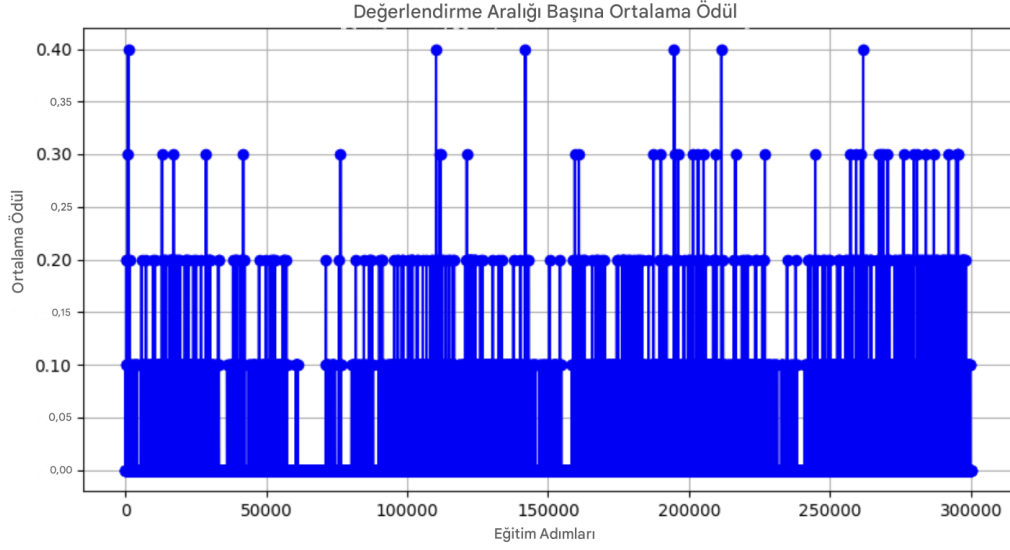
kalması nedeniyle belirli koşullar altında sınırlı bir etkiye sahip olduğunu göstermektedir.

Buna karşılık, sis faktörleri performans üzerinde daha ince etkiler göstermiştir. NUM_RED_UAV parametresi, hava ajan sayısını artırarak ortalama puanın -0.40'tan -0.41'e düşmesine neden olmuştur, bu da hafif bir negatif etki olduğunu göstermektedir. Benzer şekilde, TEAM_MEMORY ve INDIV_MEMORY'deki ayarlamalar da sınırlı değişikliklere yol açmış, bu da bu faktörlerin test edilen senaryolar içinde sonuçları etkilemedeki sınırlı rollerini yansıtmaktadır. Yer iletişimi ve iletişim mesafesi gibi iletişim parametrelerindeki değişiklikler de ortalama puanı hafifçe düşürerek -0.41'e inmiştir. Bu sonuçlar, sis faktörlerinin çevresel karmaşıklığa katkıda bulunduğunu, ancak etkilerinin ajanların kullandıkları spesifik yapılandırmalar ve stratejilere bağlı olarak değişebileceğini önermektedir. Sonuçları analiz ederken, görünürlük kısıtlamalarının (sis faktörleri) ajanların eksik bilgi nedeniyle bilinçli karar verme yeteneklerini etkilediğini, bu da daha muhafazakar stratejilere yol açtığını gözlemledik. Buna karşılık, rastgelelik (şans faktörleri), beklenmeyen olayların ortaya çıkmasına neden olarak ajanların stratejilerini hızlı bir şekilde ayarlamalarını zorlamış ve esnekliklerini test etmiştir.

6.1. Gym-Cap Optimizasyon Entegrasyonu

Yapay zekanın öğrenme ilerlemesini ve adaptasyonunu değerlendirmek için, PPO MlpPolicy algoritması kullanılarak eğitim sırasında elde edilen ortalama ödüller analiz edilmiştir. Şekil 6.4, eğitim sürecindeki öğrenme ilerlemesini, her değerlendirme aralığındaki ortalama ödülü göstermektedir. Bu grafik, ajanların zamanla uyum sağlama yeteneklerini vurgulamakta ve sistemin dinamik ve belirsiz senaryolarda ne kadar sağlam olduğunu doğrulamaktadır. Ödüllerdeki istikrarlı

artış, sis ve şans faktörlerinin entegrasyonunun yapay zeka modellerini başarıyla zorlayarak daha dayanıklı ve uyarlanabilir stratejiler geliştirmelerine olanak tanıdığını göstermektedir.



Şekil 6.4. YZ Ajanının Öğrenme İlerlemesi: PPO Eğitimi Sırasında Her Değerlendirme Aralığındaki Ortalama Ödül

6.2. Gym-Cap Sonuçları

Tablo 6.2, iki tür yapay zeka savunma ajanını, "Savunma" ajanı ve "Saldırgan" ajanını göstermektedir. Her ikisi de kırmızı takımın üsünü mavi takım saldırılarına karşı korumak için tasarlanmıştır. Savunma ajanı pasif bir strateji benimseyerek rastgele hareket desenleri ve sıfır kullanırken, Saldırgan ajanı spiral hareket desenini kullanarak mavi takım saldırganlarını aktif olarak arayıp engellemeye çalışmaktadır.

Tablo 6.3 ve Tablo 6.4, "Savunma" ve "Saldırgan" ajanlarının çeşitli mavi takım saldırganlarına karşılıklı sonuçlarını göstermektedir. Özellikle, farklı saldırı taktiklerinin savunma stratejilerinin etkinliğini nasıl değiştirdiğini vurgulamakta ve

Çizelge 6.2. YZ Savunma ve Saldırgan Ajanlarının Stratejileri

Kırmızı (Savunma)	Mavi (Saldırgan)
Savunma	Saldırgan
Rastgele	Devriye
Sıfır	Spiral
	Roomba

ajanların kontrol altında şans ve sis koşulları altında nasıl uyum sağladığını göstermektedir.

Çizelge 6.3. Savunma ve Saldırgan Ajanlarının Çeşitli Saldırı Stratejilerine Karşı Ortalama Puanları (Bölüm 1)

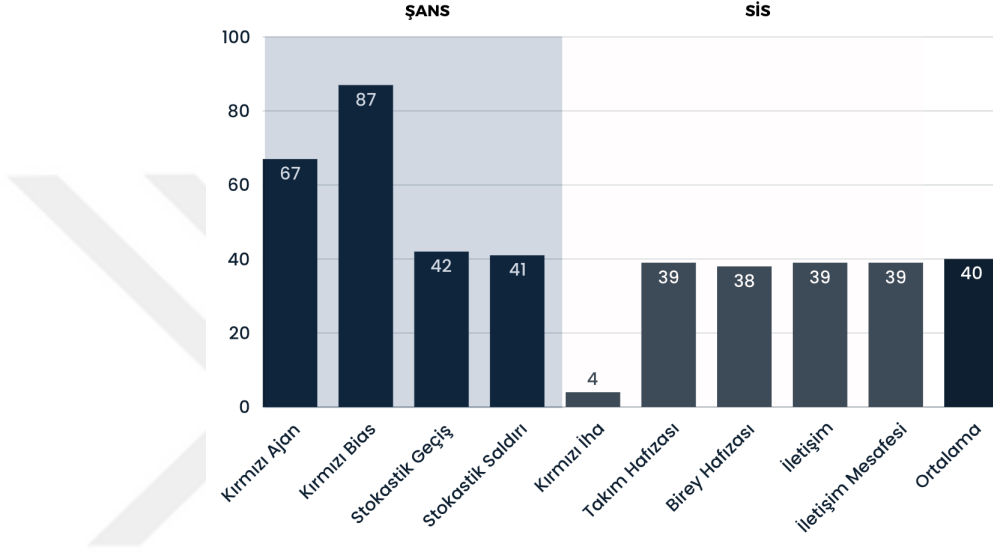
Parametreler	Savunma x Saldırgan	Rastgele x Saldırgan	Sıfır x Saldırgan	Savunma x Devriye	Rastgele x Devriye	Sıfır x Devriye	Ortalama Puan
Default	-0.95	-0.51	0.04	-1.01	-0.57	-0.06	-0.41
NUM_RED	-1.05	-0.95	-0.03	-1.10	-1.00	-0.10	-0.67
RED_ADV_BIAS	-1.07	-1.47	-0.07	-1.15	-1.38	-0.10	-0.87
STOCH_TRANSITIONS	-0.92	-0.53	0.01	-0.95	-0.54	0.04	-0.42
STOCH_ATTACK	-0.92	-0.51	-0.01	-1.02	-0.59	-0.06	-0.41
NUM_RED_UAV	-0.93	-0.58	0.03	-1.00	-0.58	-0.05	-0.40
TEAM_MEMORY	-0.93	-0.46	0.03	-1.01	-0.59	-0.06	-0.39
INDIV_MEMORY	-0.93	-0.48	-0.00	-1.00	-0.64	-0.06	-0.38
COM_GROUND	-0.92	-0.50	0.04	-1.02	-0.54	-0.06	-0.39
COM_DISTANCE	-0.94	-0.47	0.02	-0.99	-0.64	-0.07	-0.39

Çizelge 6.4. Savunma ve Saldırgan Ajanlarının Çeşitli Saldırı Stratejilerine Karşı Ortalama Puanları (Bölüm 2)

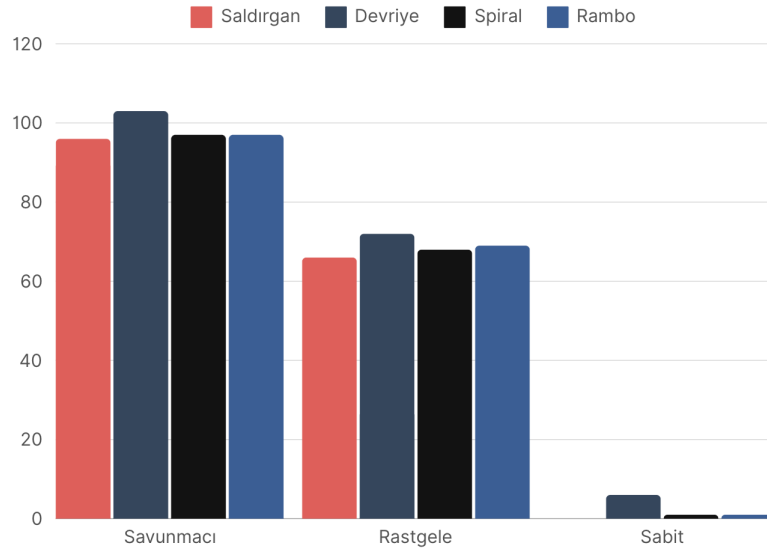
Parametreler	Savunma x Spiral	Rastgele x Spiral	Sıfır x Spiral	Savunma x Roomba	Rastgele x Roomba	Sıfır x Roomba	Ortalama Puan
Default	-0.88	-0.52	-0.01	-0.95	-0.47	0.03	-0.41
NUM_RED	-1.06	-0.92	-0.05	-1.00	-0.98	-0.06	-0.67
RED_ADV_BIAS	-1.03	-1.41	-0.01	-1.07	-1.48	-0.04	-0.87
STOCH_TRANSITIONS	-0.96	-0.61	0.02	-0.98	-0.66	0.02	-0.42
STOCH_ATTACK	-0.93	-0.53	0.02	-0.96	-0.46	0.02	-0.41
NUM_RED_UAV	-0.95	-0.52	0.01	-0.95	-0.49	0.04	-0.40
TEAM_MEMORY	-0.94	-0.57	0.01	-0.94	-0.58	0.00	-0.39
INDIV_MEMORY	-0.94	-0.48	0.01	-0.94	-0.54	0.01	-0.38
COM_GROUND	-0.94	-0.48	0.03	-0.92	-0.51	0.04	-0.39
COM_DISTANCE	-0.94	-0.60	0.03	-0.94	-0.55	0.05	-0.39

Şekiller 6.5 ve 6.6, farklı savunma stratejilerinin ortalama puanlarını görsel olarak sunmaktadır. Şekil 6.5, saldırı stratejilerinin etkisini odaklanmaktadır. Bu şekil, Saldırgan ajanının Random, Zeros ve Spiral saldırılarına karşı Savunma ajanından tutarlı bir şekilde daha yüksek ortalama puanlar elde ettiğini göstermektedir. Buna karşılık, Savunma ajanı, Roomba saldırılarına karşı Saldırgan ajanını dar bir farkla geride bırakmıştır, Savunma ajanının puanı 47 iken Saldırgan ajanının puanı 52'dir. Şekil 6.6, şans ve savaş sis koşullarının etkisini

derinlemesine incelemektedir. Özellikle, Saldırgan ajanının genel olarak Savunma ajanından daha yüksek ortalama puanlara sahip olduğu gözlemlenmiştir. Bu, Saldırgan ajanının daha saldırgan yaklaşımının belirsiz ve kısmen gözlemlenebilir ortamlara uyum sağlama konusunda daha uygun olabileceğini düşündürmektedir.



Şekil 6.5. Çeşitli Saldırı Türlerine Karşı Savunma Stratejileri İçin Ortalama Puanlar



Şekil 6.6. Ortalama Puanlar Üzerindeki Şans ve Sis Koşullarının Etkisi

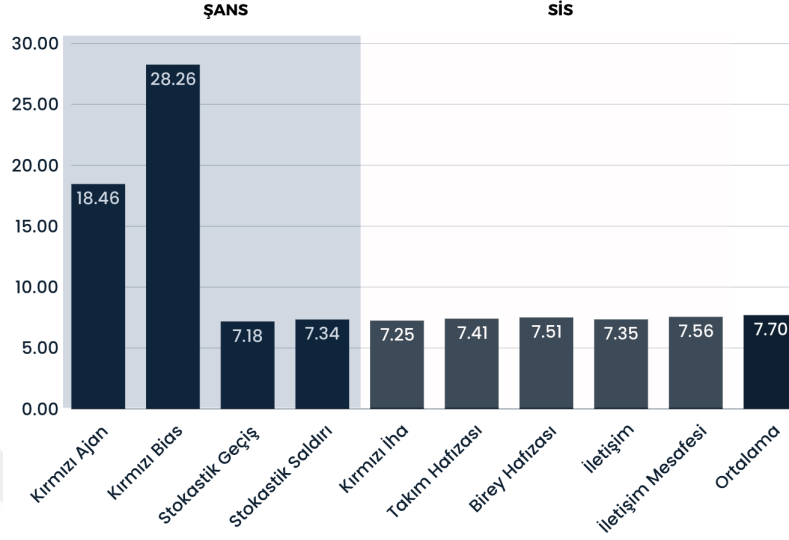
6.3. Gym-Cap Rambo Sonuçları

Mavi takım, Savunma, Devriye, Spiral, Rastgele ve Sıfır gibi çeşitli savunma stratejileri kullanmıştır. Tablo 6.5, kırmızı takımın 'Roomba' taktiğini kullanırken, mavi takımın yukarıda belirtilen savunma stratejileriyle yapılan simülasyonların sonuçlarını sunmaktadır. Kırmızı takımın 'Roomba' taktiğinin farklı şans ve sis parametreleri altında ortalama puanlarının nasıl değiştiği analiz edilmiştir. Özellikle, 'Roomba' taktiğinin çoğu senaryoda kırmızı takım için daha düşük ortalama puanlar sağladığı gözlemlenmiştir, bu da bu yaklaşımın potansiyel zayıflıklarını göstermektedir. Bu sonuçlar, stratejik esneklik eksikliği ve platform sınırlamaları nedeniyle dikkatle yorumlanmalıdır. Daha esnek bir platformun geliştirilmesi, 'Roomba' taktiği ve diğer stratejilerin daha çeşitli koşullar altında etkinliğini kesin olarak değerlendirmek için gereklidir.

Çizelge 6.5. Roomba Stratejisi ile Mavi Takım Stratejileri Arasındaki Puanlar

Parametreler	Tür	Roomba x Savunma	Roomba x Devriye	Roomba x Spiral	Roomba x Rastgele	Roomba x Sıfır	Ortalama Puan
DEFAULT	-	-7.56	-7.30	-8.38	-8.46	-7.96	-7.73
NUM_RED	Şans	-18.50	-17.19	-18.57	-17.99	-18.04	-18.46
RED_ADV_BIAS	Şans	-30.23	-23.88	-28.21	-28.56	-29.42	-28.26
STOCH_TRANSITIONS	Şans	-6.80	-7.78	-7.10	-7.37	-6.84	-7.18
STOCH_ATTACK	Şans	-7.43	-6.58	-7.49	-7.48	-7.71	-7.34
NUM_RED_UAV	Sis	-7.17	-6.96	-7.64	-8.42	-8.07	-7.25
TEAM_MEMORY	Sis	-7.47	-7.92	-7.03	-7.01	-7.61	-7.41
INDIV_MEMORY	Sis	-7.52	-7.92	-7.26	-7.24	-7.59	-7.51
COM_GROUND	Sis	-7.06	-6.85	-7.09	-7.12	-7.61	-7.35
COM_DISTANCE	Sis	-7.52	-7.78	-7.29	-7.76	-7.47	-7.56

Şekil 6.7, her strateji için ortalama puanları göstermektedir; şans faktörleri sol tarafta ve sis faktörleri sağ taraftadır. Parametre değişimlerinde, kırmızı takımın Roomba stratejisi için ortalama puanları ile , mavi takımın bir kaç Savunma stratejisi için ortalama puanları temsil etmektedir.



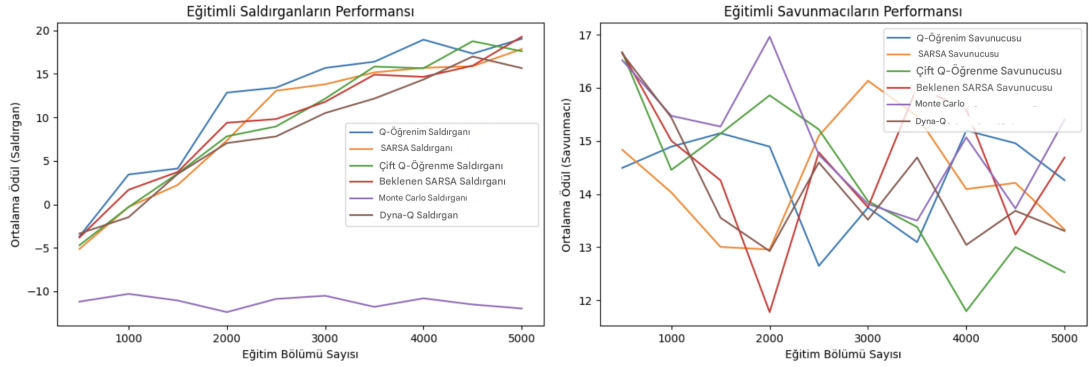
Şekil 6.7. Farklı Koşullar Altında Roomba Stratejisi İçin Ortalama Puanlar

6.4. Pekiştirmeli Modellerin Karşılaştırma Sonuçları

Şekil 6.8, çalışmada kullanılan farklı yapay zeka modellerinin (Q-Öğrenme, SARSA, Double Q-Öğrenme, Expected SARSA, Monte Carlo ve Dyna-Q) performanslarını karşılaştırmaktadır. Bu grafiklerde, modellerin saldırgan ve savunmacı rollerindeki başarıları, CPU ve RAM kullanımları, toplam eğitim süresi ve adım kullanımı üzerinden analiz edilmiştir.

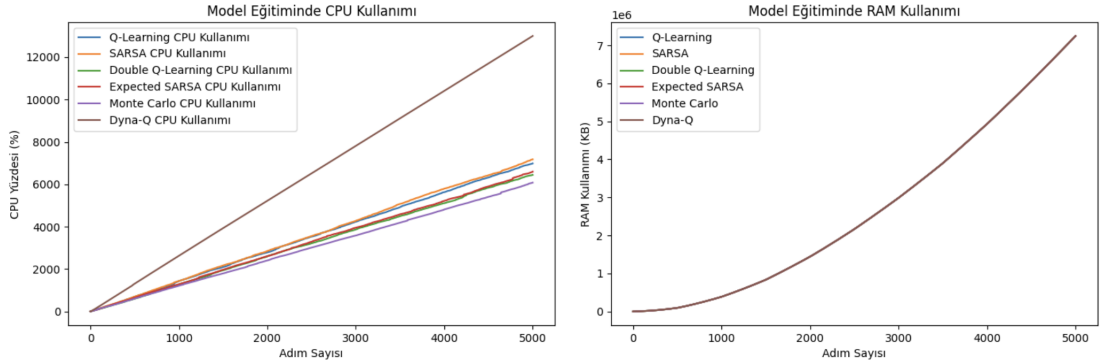
Şekil 6.8 grafiklerde, saldırgan ve savunmacı rollerindeki ortalama ödüllerin eğitim bölümleri boyunca nasıl değiştiğini göstermektedir. Saldırgan performansında, Q-Öğrenme algoritması diğer modellere kıyasla tutarlı bir artış göstererek en başarılı sonuçları elde etmiştir. Bu durum, Q-Öğrenme'in stratejik karar verme süreçlerini daha etkili bir şekilde optimize ettiğini ortaya koymaktadır. Ancak savunmacı rolündeki performans artışı tüm modellerde sınırlı kalmıştır. Bunun nedeni, savunmacının hedefinin (saldırganın) sürekli hareketli olması ve ajanların

bu duruma uyum sağlayacak şekilde yeterince eğitilememiş olmasından kaynaklanabilir.



Şekil 6.8. Farklı Modellerin Sonuçlarının Karşılaştırması

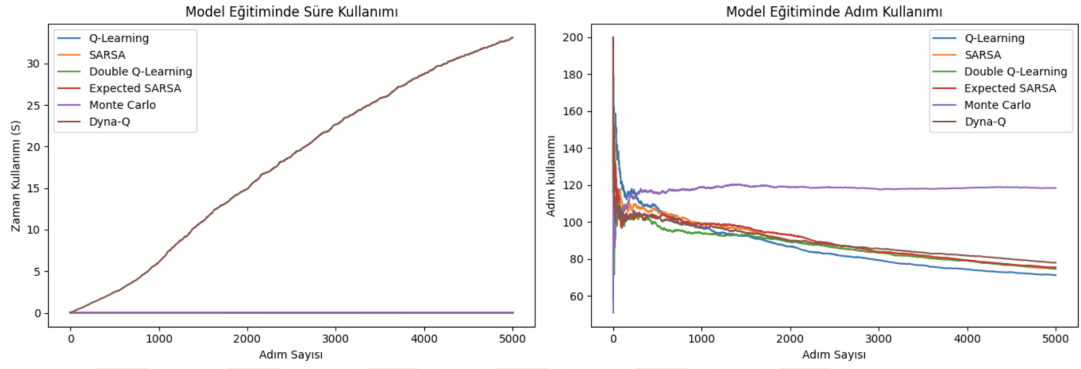
Şekil 6.9 grafikler, CPU ve RAM kullanımlarını analiz etmektedir. Eğitim süresince RAM kullanımı tüm modellerde kademeli olarak artmaktadır, çok ufak değişimler yüzünden grafik çizgisi tek olarak görünmektedir. CPU kullanımı açısından bakıldığında, genelde benzer tüketimler görünürken, Dyna-Q diğerlerinden biraz farkla daha fazla tüketim gerçekleştirilmiştir. En düşük Monte Carlo olurken, SARSA ve Q-Öğrenme modelleri ortalama bir tüketim sergilemiştir.



Şekil 6.9. Farklı Modellerin CPU-RAM Karşılaştırması

Şekil 6.9 grafikler, model eğitiminde geçen toplam süreyi göstermektedir. Bu metrik, modellerin hesaplama maliyetlerini değerlendirirken kritik bir rol

oyunmaktadır. Burada Q-Öğrenme algoritmasının nispeten daha kısa sürede başarılı sonuçlar ürettiği ve hesaplama açısından en verimli yöntemlerden biri olduğu gözlemlenmiştir. Son grafikte ise modellerin adım başına gerçekleştirdiği işlemler karşılaştırılmıştır. Q-Öğrenme, adım kullanımında da optimize bir performans sergileyerek hem saldırgan hem de savunmacı rollerinde daha az adımda daha yüksek başarı sağlamıştır.



Şekil 6.10. Farklı Modellerin Zaman ve Adım Karşılaştırması

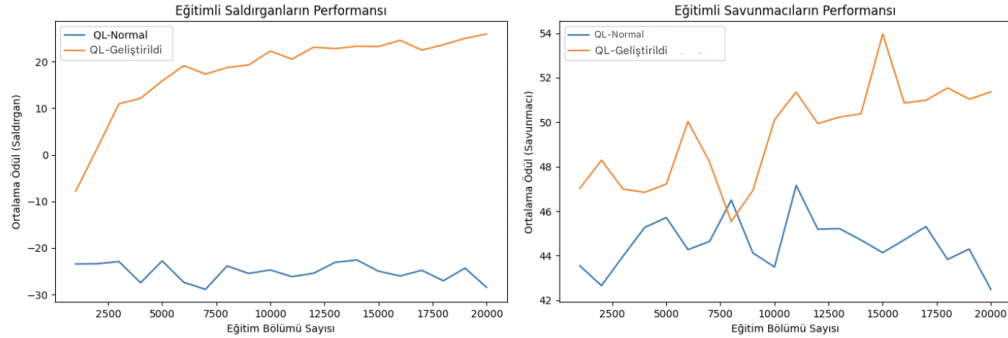
Sonuç olarak, bu grafikler Q-Öğrenme'nin diğer modellere göre hem kaynak verimliliği hem de performans açısından üstün olduğunu ortaya koymaktadır. Bu nedenle, Q-Öğrenme çalışmada tercih edilen ana model olarak seçilmiştir. Bununla birlikte, savunmacı performansındaki sınırlı artışın, hedefin hareketliliği ve eğitim süreçlerinin bu durum için yeterince optimize edilmemesi ile ilişkili olduğu değerlendirilmiştir. Gelecekteki çalışmalarda, dinamik hedef davranışlarını daha iyi modellemek ve optimize etmek için ek yöntemler araştırılabilir.

6.5. Q-öğrenme Model Karşılaştırma Sonuçları

Şekil 6.11, normal Q-Öğrenme modeli (QL-Normal) ile otomatik parametre optimizasyonuna sahip geliştirilmiş Q-Öğrenme modeli (QL-Developed) arasındaki

performans farklarını göstermektedir. Eğitimli saldırganların performansında (sol grafik), QL-Developed modeli, tüm eğitim bölümleri boyunca QL-Normal modeline kıyasla belirgin bir üstünlük sergilemiştir. Bu üstünlük, QL-Developed modelinde kullanılan otomatik parametre optimizasyonu sayesinde ajanların hem keşif (ϵ) hem de öğrenme hızı (α) değerlerini dinamik olarak ayarlayabilmesi ile ilişkilidir. Bu adaptasyon, ajanların stratejik karar alma süreçlerini iyileştirmiş ve daha hızlı öğrenmelerini sağlamıştır.

Savunmacı performansında (sağ grafik) ise her iki model arasında daha sınırlı bir fark görülmüştür. Bununla birlikte, QL-Developed modelinin savunmacı rolünde de genel olarak daha yüksek ortalama ödül elde ettiği gözlemlenmiştir. Bu durum, savunmacı ajanın hedefinin (saldırgan) hareketli olmasından dolayı eğitim süreçlerinde genel olarak sınırlı bir başarı sergilemesine rağmen, otomatik parametre optimizasyonunun bu sınırlamayı kısmen de olsa telafi ettiğini göstermektedir. Sonuç olarak, QL-Developed modeli, hem saldırgan hem de savunmacı rollerinde daha etkili bir performans sergileyerek çalışmanın en başarılı modeli olarak öne çıkmıştır.



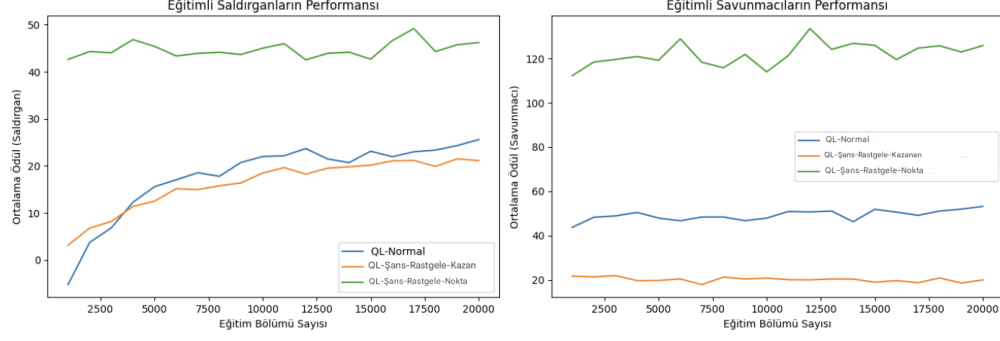
Şekil 6.11. QL-Normal ve QL-Developed Modellerinin Performans Karşılaştırması

6.6. Şans Sonuçları

Şekil 6.12, geliştirilmiş Q-Öğrenme modeli (QL-Normal) ile şans faktörlerini içeren Q-Luck-Random-Win ve Q-Luck-Random-Point modellerinin performans karşılaştırmasını göstermektedir. Sol grafikte, Saldırgan ajanların performansları karşılaştırılmıştır. Q-Luck-Random-Point modeli, rastgele ek puan kazanma mekanizmasının etkisiyle saldırgan performansında genel olarak diğer modellerin üzerinde bir başarı göstermiştir. Bu durum, rastgele ödüllerin strateji dışı avantajlar sunarak saldırganların ortalama ödülleri artırdığını göstermektedir. Ancak Q-Luck-Random-Win modeli, rastgele kazanma mekanizmasının düşük olasılıkla gerçekleşmesi ve bu mekanizmanın ajanların öğrenme sürecine daha sınırlı bir katkı sağlamasından dolayı QL-Normal modelinin altında bir performans sergilemiştir.

Sağ grafikte, savunmacı ajanların performansları değerlendirilmiştir. Q-Luck-Random-Point modeli, rastgele ek puan mekanizması sayesinde savunmacı rolünde de en yüksek ortalama ödülü elde etmiştir. Ancak, bu durum savunmacı stratejilerin etkili bir şekilde optimize edilmesi yerine, rastgele puanların performansı doğrudan artırmasından kaynaklanmaktadır. Q-Luck-Random-Win modeli ise, rastgele kazanma durumunun düşük sıklıkla gerçekleşmesi nedeniyle savunmacı performansında en düşük başarıyı göstermiştir. QL-Normal modelinin ise, şans faktörlerinden etkilenmeyen daha istikrarlı bir performans sergilediği görülmüştür.

Genel olarak, bu sonuçlar şans faktörlerinin oyun ortamında ödüller ve stratejik performans üzerindeki doğrudan etkisini vurgulamaktadır. Özellikle, Q-Luck-Random-Point modeli, rastgele ek puan mekanizmasının hem saldırgan hem de savunmacı rollerindeki genel performansı artırdığını göstermiştir.



Şekil 6.12. QL-Normal ve Şans Faktörleri İçeren Modellerin Performans Karşılaştırması

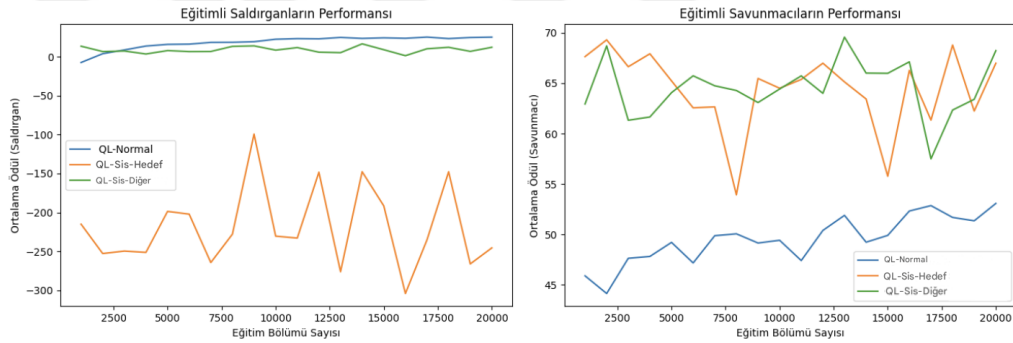
6.7. Sis Sonuçları

Şekil 6.13, geliştirilmiş Q-Öğrenme modeli (QL-Normal) ile sis faktörlerini içeren QL-Fog-Target ve QL-Fog-Other modellerinin performans karşılaştırmasını göstermektedir. Sol grafikte, saldırganların performansına bakıldığında QL-Normal modeli, ajanların çevresel tüm bilgiyi görebilmesi nedeniyle en yüksek performansı sergilemiştir. QL-Fog-Other modeli, ajanların yalnızca kendi konumlarını ve hedeflerinin konumlarını görebilmesi durumunda, QL-Normal modeline kıyasla daha düşük bir performans göstermiştir. QL-Fog-Target modeli ise, ajanların sadece kendi konumlarını görebilmesi nedeniyle, saldırgan performansında belirgin bir şekilde düşük kalmıştır. Bu sonuçlar, bilgi eksikliğinin stratejik karar alma süreçlerini ve saldırganların genel ödül kazanımını önemli ölçüde etkilediğini ortaya koymaktadır.

Sağ grafikte, savunmacı ajanların performansları değerlendirilmiştir. QL-Fog-Other modeli, savunmacıların hem kendi hem de hedeflerinin (saldırgan) konumlarını görebildiği durumlarda, savunma performansında en yüksek ortalama ödülü elde etmiştir. QL-Fog-Target modeli, sadece kendi konumlarını görebilen savunmacıların stratejik hareketlerinin daha sınırlı olması nedeniyle daha düşük bir performans

sergilemiştir. QL-Normal modeli ise tüm bilgiyi görebilmesine rağmen, savunmacıların hedeflerinin hareketli olması ve dinamik koşulların optimize edilememesi nedeniyle diğer modellerle karşılaştırıldığında beklenen üstünlüğü sağlayamamıştır.

Bu sonuçlar, ajanların bilgi seviyelerinin performanslarını doğrudan etkilediğini göstermektedir. Tam bilgiye dayalı stratejilerin avantaj sağladığı görülmekle birlikte, kısmi bilginin (QL-Fog-Other) belirli durumlarda daha stratejik hareketler yapılmasına olanak tanıdığı, özellikle savunmacı rolünde önemli bir katkı sunduğu anlaşılmıştır.



Şekil 6.13. QL-Normal ve Sis Faktörleri İçeren Modellerin Performans Karşılaştırması

Sonuç olarak; bulgularımız yapay zeka destekli stratejik savaş oyunlarında gerçekçiliği artırma yönünde önemli bir adım olduğunu göstermektedir. Yapay zeka tarafından eğitilen sistemlerin, savaş sisini (belirsizlik) ve şansını (rastgelelik) entegre ederek dinamik ve öngörülemez ortamları simüle edebilme potansiyeli gösterilmiştir. Bulgular, bu faktörlerin oyun sonuçları üzerindeki önemli etkilerini vurgulamakta ve belirsizlik altında karar verme süreçlerini iyileştirmek için yapay zeka eğitim modellerine entegre edilmelerinin önemini göstermektedir. Bu çalışma, stratejik simülasyonlara derinlik katma potansiyelini vurgulayarak, bunların sadece daha gerçekçi değil, aynı zamanda daha eğitici hale gelmesini sağlamaktadır.

7. TARTIŞMA

7.1. Araştırma Sorusu Üzerine Bulgular

Araştırma sorumuzu yanıtlamada, savaş alanı ortamının çok yönlü doğasına dair bulgularımız, oyun sonuçlarının şans ve savaş sisteminin unsurları tarafından nasıl derinden etkilendiğini göstermektedir. Bu bulgular, belirsizlik ve yapay zeka kullanımının stratejik karar alma süreçlerini nasıl etkilediğine dair geniş anlayışla uyumludur, bu sadece oyun oynamada değil, aynı zamanda uluslararası ilişkiler ve stratejik güven konularında da geçerlidir (2023). Özellikle, ajan sayısındaki ve önyargılardaki (örneğin, karakterlerin güçlendirilmesi) varyansın etkisi, diğer faktörlere kıyasla daha belirgin olmuştur. Bu gözlem, oyun ortamındaki karmaşıklığı vurgulamakta ve stokastik unsurların genel sonuçları belirlemedeki kritik rolünü pekiştirmektedir, bu da stratejik oyun dinamikleri ve yapay zekanın adaptasyon yetenekleri üzerine geniş tartışmalarımızı beslemektedir.

Ayrıca, çalışmamızın "NUM_RED Chance" ve "STOCH_ATTACK Chance" parametrelerini incelemesi, özellikle Roomba x Savunma ve Roomba x Devriye stratejileri içinde farklı stratejik sonuçlar üzerinde önemli bir etki yaratmıştır. Bu tür ayrımlar, şans ve stratejik önyargının çeşitli oyun senaryolarında nasıl nüanslı etkiler yarattığını vurgulamakta, yapay zekanın stratejik karar alma ve adaptasyon süreçlerine olan etkisini güçlendirmektedir. "COM_GROUND Sis" parametresinin araştırılması ise strateji etkinliğinde çevresel değişkenlerin farklı etkilerini açıklığa kavuşturarak, stratejik oyun senaryolarında çevresel değişkenlerin stratejik sonuçları ve yapay zeka performansı üzerindeki farklı etkilerini ortaya koymaktadır. Bu bulguları sentezleyerek, çalışmamız yapay zeka modelleri ile stratejik savaş simülasyonları arasındaki karmaşık etkileşime dair yenilikçi içgörüler sunmaktadır.

Her bir parametreyi izole ederek, sis ve şans faktörlerinin stratejik karar alma süreçlerine bireysel katkılarını belirleyebilmiş ve sonuçlarımızın netliğini artırmış bulunmaktayız. Şans ve çevresel faktörlerin önemli etkilerini vurgulayarak, stratejik oyunların dinamiklerinde yapay zeka sistemlerinin nasıl uyum sağlayabileceğini ve belirsizlikle başa çıkma yeteneklerini geliştirebileceğini göstermekteyiz. Bu araştırma, gelecekteki stratejik planlama ve karar alma süreçlerinde yapay zekanın rolünü derinleştirmek için bir temel oluşturmaktadır.

7.2. Çalışmanın Sınırlamaları

Araştırma sürecimiz boyunca "gym-cap" kütüphanesi ile ilgili birkaç zorlukla karşılaştık. Paket, uzun süre önce geliştirilmiş olduğu için daha yeni Python paketleriyle uyumluluk sorunları yaşanmış ve Gym-Cap platformu artık aktif olarak desteklenmemektedir. Bu bakım eksikliği, kırmızı takım için saldırgan benzeri stratejilerin uygulanmasını zorlaştırmış ve orijinal geliştiricilere ulaşma imkanımızı engellemiştir. Sonuç olarak, RAIDE projesinin özelleştirilmiş CTF paketine güvenmek zorunda kaldık; bu paket, Gym gibi eski çerçevelerden Gymnasium'a geçiş yaparak daha sağlam bir ortam sunmuştur. Bu ayarlamaya rağmen, platformun esnek olmaması, gelişmiş taktiksel manevraların ve koordine takım hareketlerinin uygulanmasını sınırlamıştır. Örneğin, kısıtlı eylem alanı, ajanların karmaşık hareket desenleri veya işbirlikçi saldırgan taktikler yürütmesini engellemiş, bu nedenle Kırmızı takım için 'Roomba' taktiğinin benimsenmesine ve Mavi takım için savunma stratejilerinin kullanılmasına neden olmuştur. Bu değişiklikler, Mavi takımın saldırgan etkinliğini önemli ölçüde azaltmış, bu da sonuçların doğruluğunu ve genellenebilirliğini etkilemiştir. Bu durum, yapay zeka destekli stratejik simülasyonların kapsamlı test edilmesi için uygun simülasyon

platformlarının seçilmesi ve bakımı gerekliliğini ortaya koymaktadır. Gym-Cap'in kısıtlamalarını aşmak amacıyla bu çalışmanın ikinci aşamasında daha temel düzeyde bir CTF uygulaması geliştirilmiştir; ancak bu uygulamada da parametrelerin (öğrenme oranı, keşif oranı, vb.) en uygun değerlerinin belirlenmesi zor olmuştur. Gelecekteki araştırmalar, daha esnek platformlara geçiş yapmayı veya mevcut olanları daha gelişmiş strateji uygulamalarını desteklemek için özelleştirmeyi düşünmelidir, böylece yapay zeka tabanlı saldırgan ve savunma stratejilerinin sağlamlığını ve uygulanabilirliğini artırmış oluruz.

7.3. Gelecek Araştırmalar

Bu araştırma, gelecekteki çalışmalara temel oluşturmakta olup, optimize edilmiş stratejilerin geliştirilmesi ve öğrenme hızlarını artırmak için daha basit CTF ortamlarının keşfedilmesi üzerine odaklanacaktır. Bu yaklaşım, uyum sağlama ve strateji geliştirme yeteneklerini değerlendirmek için temel bir platform oluşturmayı amaçlamaktadır. Temel hedeflerden biri, otonom saldırı sistemlerinin dinamik olarak taktiklerini yeniden kalibre edebilecek şekilde geliştirilmesi olup, bu da stratejik oyunlardaki denge ve rekabeti artırmayı hedeflemektedir. Gelecek çalışmalarımızda, halihazırda geliştirilmiş olan temel düzey CTF uygulamasını daha da iyileştirerek ödül ve ceza mekanizmalarını daha yalın bir yapıya kavuşturmayı ve böylece öğrenmenin daha doğrudan gözlemlenmesini sağlamayı amaçlamaktayız. Bu sayede, sis ve şans faktörlerinin eğitim üzerindeki etkileri daha etkin şekilde incelenebilirken, parametre optimizasyonu konusundaki zorluklar da daha iyi anlaşılacaktır. Bu yaklaşım, sis ve şans faktörlerinin eğitim üzerindeki etkilerini daha etkili bir şekilde gözlemlememize olanak tanıyacaktır. Daha sonra, çeşitli platformlar üzerinde yapılacak deneyler, farklı pekiştirmeli öğrenme

metodolojileri aracılıđıyla üstün sonuçların elde edilmesini sađlayacak farklılıkları ortaya koyacaktır.

Geleneksel HRL (Hiyerarşik Pekiştirmeli Öğrenme) yaklaşımlarından sapmanın yanı sıra, takım içi ve bireysel belleklerin oyun içi oynama üzerindeki etkilerini daha geniş kapsamlı olarak incelemeyi ve düşmanca yapay zeka yapıları ile etkileşime giren katılımcıların öğrenme hızlarını değerlendirmeyi amaçlamaktayız. Ayrıca, oyun mekaniğini geliştirmek için takım koordinasyonunu daha gerçekçi bir şekilde simüle edebilen yeni özellikler ekleyerek stratejik derinliği ve gerçekçiliđi artırmayı planlamaktayız. Bu tür bir araştırma, bilişsel dinamiklerin stratejik karar alma ve uyum sağlama süreçlerini oyun ortamlarında nasıl etkilediğini aydınlatacaktır. Bu çabalar, stratejik oyunlardaki rekabetçi adilliđi artırmayı ve kompleks stratejik planlama ve yürütme konularında yapay zekanın rolünü ilerletmeyi hedeflemektedir.

8. SONUÇ

Bu çalışma, yapay zeka destekli stratejik savaş oyunlarında gerçekçiliği artırmaya yönelik önemli bir adım atarak, sis ve şans faktörlerinin etkisini anlamaya odaklanmıştır. Elde edilen sonuçlar, bu faktörlerin stratejik düşünme, karar alma ve öğrenme süreçleri üzerindeki etkilerini çarpıcı bir şekilde ortaya koymuştur. Özellikle, temel CTF oyunu ile yapılan testler, sis ve şans faktörlerinin oyuncuların ve yapay zeka ajanlarının stratejik performansını anlamlı ölçüde değiştirdiğini göstermiştir. Sis ve şans faktörlerinin, belirsizlik altında öğrenme ve adaptasyon süreçlerine nasıl farklı etkiler yaptığı detaylı bir şekilde sınıflandırılmıştır.

Bu bağlamda yapılan analizler, sis faktörünün genellikle bilgi kısıtlaması yoluyla karar verme süreçlerini zorlaştırdığını ve oyuncuların belirsizlik yönetimi becerilerini geliştirdiğini ortaya koyarken, şans faktörünün rastlantısal olaylar aracılığıyla oyun dinamiklerine beklenmedik değişiklikler kattığı gözlemlenmiştir. Veriler, sis ve şansın birbirinden bağımsız olarak öğrenme üzerinde farklı etkiler yarattığını, her bir parametre değişikliğinin oyuncuların ve yapay zeka ajanlarının stratejik öğrenme süreçlerine farklı şekillerde katkıda bulunduğunu göstermektedir. Özellikle, CTF oyunlarında yapılan deneysel çalışmalarda, sis faktörünün bilgi eksikliği ile tetiklenen daha dikkatli ve ölçülü stratejilere yol açarken, şans faktörünün adaptasyon ve hızlı karar alma yeteneğini ön plana çıkardığı tespit edilmiştir.

Araştırma ayrıca, sis ve şans faktörlerinin yapay zeka tabanlı simülasyonların gerçekçiliğini artırmakla kalmayıp, strateji geliştirme, hızlı uyum sağlama ve belirsizlik yönetimi gibi becerileri geliştirdiğini göstermektedir. Örneğin, yapılan sınıflandırmalar, hangi oyun dinamiklerinin sis veya şans faktörleri altında nasıl

değiştirdiğini netleştirerek, bu faktörlerin modelleme ve simülasyon süreçlerindeki önemini daha somut hale getirmiştir. Bu tür bir ayrıştırma, stratejik simülasyonlarda öğrenme süreçlerinin daha etkin bir şekilde yönlendirilmesine olanak sağlamaktadır.

Bu çalışma yapay zeka destekli stratejik simülasyonlarda gerçekçiliği artırmak ve stratejik düşünme becerilerini geliştirmek adına önemli bir zemin hazırlamaktadır. Sis ve şans faktörlerinin strateji geliştirme süreçlerine olan etkileri sistematik bir şekilde ele alınmıştır. Bu bulgular, stratejik simülasyonların sadece akademik araştırmalarda değil, gerçek dünya uygulamalarında da daha etkin bir şekilde kullanılmasını sağlama potansiyeli taşımaktadır.

Elde edilen bulgular, askeri eğitimden siber güvenlik tatbikatlarına, afet yönetiminden ticari uygulamalara kadar geniş bir alanda uygulanabilir. Özellikle askeri eğitimlerde, sis ve şans faktörlerinin yer aldığı oyunların, personelin risk analizi, belirsizlik yönetimi ve strateji oluşturma becerilerini geliştirmektedir. Siber güvenlik bağlamında, rastlantısal saldırı senaryolarının simülasyonu, savunma mekanizmalarının güçlendirilmesini sağlamaktadır. Afet yönetiminde ise belirsiz ve kısmen rastlantısal gelişen durumların modellenmesi, müdahale ekiplerinin karar verme yetkinliklerini artırmaktadır.

Gelecekteki araştırmalar, sis ve şans faktörlerinin daha karmaşık dinamiklerdeki etkilerini incelemeli ve iletişim kısıtları, takım koordinasyonu ve bellek kapasitesi gibi ek parametrelerin strateji oluşumuna katkılarını değerlendirmelidir. Bu tür araştırmalar, yapay zeka ajanlarının belirsizlik altında daha esnek ve öğrenen sistemler geliştirmesine olanak sağlayacaktır. Ayrıca, sis ve şans faktörlerinin oyun içi karar verme ve takım içi rol dağılımına etkilerini daha derinlemesine incelemek, bu alanda yeni katkılar sunacaktır.

KAYNAKÇA

- Adil, K., Jiang, F., Liu, S., Jifara, W., Tian, Z., & Fu, Y. (2017). Real-time strategy (rts) games—artificial intelligence and starcraft: Current state and open challenges. *International Journal of Advanced Computer Science & Applications*, 8(12), 16–24.
- Balogh, Á., Érsök, M., Erdődi, L., Szarvák, A., Kail, E., & Bánáti, A. (2022, Mar). Ctf game optimization based on sql injection attacks. In *2022 ieee 20th jubilee world symposium on applied machine intelligence and informatics (sami)* (pp. 000153–000158). IEEE.
- Beason, J., Novitzky, M., Kliem, J., Errico, T., Serlin, Z., Becker, K., ... others (2024). Evaluating collaborative autonomy in opposed environments using maritime capture-the-flag competitions. *arXiv preprint arXiv:2404.17038*.
- Cai, C., & Zhang, R. (2023). Malicious use of ai, uncertainty, and us–china strategic mutual trust. In *The palgrave handbook of malicious use of ai and psychological security* (pp. 377–396). Cham: Springer International Publishing.
- Chetwyn, R. A., & Erdődi, L. (2022, Dec). Towards dynamic capture-the-flag training environments for adversarial security agents with reinforcement learning. In *2022 ieee international conference on big data (big data)* (pp. 2585–2594). IEEE.
- Cole, S. V. (2022, Jul). Impact of capture the flag (ctf)-style vs. traditional exercises in an introductory computer security class. In *Proceedings of the 27th acm conference on innovation and technology in computer science education vol. 1* (pp. 470–476).
- Dachowicz, A., Mall, K., Balasubramani, P., Maheshwari, A., Raz, A. K., Panchal, J. H., & DeLaurentis, D. A. (2022). Mission engineering and design using

- real-time strategy games: An explainable ai approach.
- Doe, J., & Lee, S. (2023). Comparative analysis of reinforcement learning techniques in dynamic strategy games. *Journal of Artificial Intelligence Research*.
- Erdődi, L., & Zennaro, F. M. (2022). Modeling web hacking with reinforcement learning: Agent web model. *International Journal of Information Security*, 21(2), 293–309.
- Farshi, E. (2023). Streak clustering in sports: Applications and impact on performance and strategy.
- Frank, A. B. (2022). Artificial intelligence without ai. *The Journal of Defense Modeling and Simulation*, 15485129221074352.
- Goecks, V. G., Waytowich, N., Asher, D. E., Jun Park, S., Mittrick, M., Richardson, J., & Kott, A. (2023). Using games and simulators as a platform for developing ai for command and control.
- Harris, B. N., & Freeman, S. (2023). Bridging a virtual gap: Using wargaming as a remote teaching tool. *PS: Political Science & Politics*, 56(3), 431–437.
- Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., & Graepel, T. (2019a). Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science*, 364(6443), 859–865.
- Jaderberg, M., Czarnecki, W. M., Dunning, I., Marris, L., Lever, G., Castaneda, A. G., ... others (2019b). Human-level performance in 3d multiplayer games with population-based reinforcement learning. *Science*, 364(6443), 859–865.
- Kliem, J., & Dasgupta, P. (2023). Reward shaping for improved learning in real-time strategy game play. *arXiv preprint arXiv:2311.16339*.
- Kucek, S., & Leitner, M. (2020). An empirical survey on functions and configurations of open-source capture the flag (ctf) environments. *Journal of Network and*

Computer Applications, 151, 102470.

- Minagar, S., & Sakzad, A. (2023, Jun). Automatic problem generation for ctf-based changes in it forensics courses. In *Proceedings of the 2023 conference on innovation and technology in computer science education v. 1* (pp. 229–235).
- Murray, N. A. (2023). Where intelligences dare to gamble: Clausewitz’s card game analogies. *The US Army War College Quarterly: Parameters*, 53(2), 12.
- Obeia, K. O., & Wójcik, A. (2023). Comparing ‘fog of war’ models in digital warfare games: An entity-component-system architecture and arcgis approach.
- Savin, G. M., Asseri, A., Dykstra, J., Goohs, J., Melaragno, A., & Casey, W. (2023, Aug). Battle ground: Data collection and labeling of human cyber operators through ctf games. In *Proceedings of the 16th cyber security experimentation and test workshop* (pp. 32–40).
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419), 1140–1144.
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., & Hassabis, D. (2017). Mastering the game of go without human knowledge. *Nature*, 550(7676), 354–359.
- Smith, R., & Tanaka, Y. (2024). Heuristic learning strategies in multi-agent competitive games. *IEEE Transactions on Neural Networks and Learning Systems*.
- Spencer, P., Dasgupta, P., McCarrick, M., Novitzky, M., Hubczenko, D., Redfield, S., & Mittu, R. (2021). Increasing resilience under adversarial conditions through ai-supported strategic gameplay.
- Van Stolen, N., Kim, S. H., Tran, H. T., & Chowdhary, G. (2020, May). Assessing the adaptation performance of hierarchical deep reinforcement learning. In

2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 11457–11463). IEEE.

Vinyals, O., Babuschkin, I., Chung, J., Mathieu, M., Jaderberg, M., Czarnecki, W. M., & Silver, D. (2019). *Alphastar: Mastering the real-time strategy game starcraft ii* (Vol. 2). DeepMind blog.

Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dykstra, A., Chung, J., & Silver, D. (2019). Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782), 350–354.

Zennaro, F. M., & Erdődi, L. (2023). Modeling penetration testing with reinforcement learning using capture-the-flag challenges: Trade-offs between model-free learning and a priori knowledge. *IET Information Security*.

Zhong, L. (2024). Comparison of q-learning and sarsa reinforcement learning models on cliff walking problem. In *Proceedings of the 2023 international conference on data science, advanced algorithm and intelligent computing (dai 2023)*.

ÖZGEÇMİŞ

Y**** Ö***

KİŞİSEL BİLGİLER

Doğum Yeri : Ö****
Doğum Yılı : 1***
Uyruğu : T*
Adres : T***** M*** K***** * S**** N* ** K***** M****
E-posta : y*****@g*****

ARAŞTIRMACI BİLGİLERİ

Öğrenci Orcid ID : 0009-0005-9989-4343
Danışman Orcid ID : 0000-0002-3595-8783

EĞİTİM BİLGİLERİ

Lisans : Muğla Sıtkı Koçman Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendisliği Bölümü, 2017
Staj : UMEA Üniversitesi, Teknoloji Fakültesi, Bilgisayar Bilimleri
Bölümü, 2023

İŞ DENEYİMİ

2023 – : Co-Founder CTO: MediTechLabs Yazılım
2022 – 2023 : Founder - Developer: TAGITECH
2021 – 2022 : AI - Web Developer: Ekol Hospitals
2021 – 2021 : Senior Full Stack Developer: CoreBit
2018 – 2022 : Mid-level Full Stack Developer: Idealtur
2017 – 2018 : Junior Software Engineer: Akım Tekno