



**T.C.**  
**KONYA TEKNİK ÜNİVERSİTESİ**  
**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**



**HAPS-DESTEKLİ V2X SİSTEMLERDE DERİN  
ÖĞRENME TABANLI VE BİLGİ YAŞINA  
DUYARLI KAYNAK YÖNETİMİ**

**Ahmet Melih İNCE**

**YÜKSEK LİSANS TEZİ**

**Elektrik-Elektronik Mühendisliği Anabilim Dalı**

**Haziran-2025**  
**KONYA**  
**Her Hakkı Saklıdır**

## TEZ KABUL VE ONAYI

Ahmet Melih İNCE tarafından hazırlanan “HAPS-DESTEKLİ V2X SİSTEMLERDE DERİN ÖĞRENME TABANLI VE BİLGİ YAŞINA DUYARLI KAYNAK YÖNETİMİ” adlı tez çalışması .../.../... tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

### Jüri Üyeleri

#### Başkan

Doç. Dr. Akif DURDU  
Konya Teknik Üniversitesi

#### Danışman

Doç. Dr. Ayşe Elif CANBİLEN  
Konya Teknik Üniversitesi

#### Üye

Doç. Dr. Hakkı SOY  
Necmettin Erbakan Üniversitesi

### İmza

.....

.....

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mevlüt UYAN  
Enstitü Müdürü

## **TEZ BİLDİRİMİ**

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

## **DECLARATION PAGE**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Ahmet Melih İNCE

Tarih:

# ÖZET

## YÜKSEK LİSANS TEZİ

### HAPS-DESTEKLİ V2X SİSTEMLERDE DERİN ÖĞRENME VE BİLGİ YAŞINA DUYARLI KAYNAK YÖNETİMİ

Ahmet Melih İNCE

Konya Teknik Üniversitesi  
Lisansüstü Eğitim Enstitüsü  
Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Ayşe Elif CANBİLEN

2025, 64 Sayfa

Jüri

Doç. Dr. Ayşe Elif CANBİLEN  
Doç. Dr. Akif DURDU  
Doç. Dr. Hakkı SOY

Bu tez çalışmasında; yüksek irtifa platform istasyonu (high altitude platform station, HAPS) destekli araçtan-her şeye (vehicle-to-everything, V2X) haberleşme sistemleri için çok ajanlı derin deterministik politika gradyanı (multi agent deep deterministic policy gradient, MADDPG) algoritmasına dayalı yeni bir kaynak yönetim çerçevesi önerilmektedir. Geliştirilen sistem; bilgi güncelliği, enerji verimliliği ve iletişim kapasitesi gibi temel performans göstergelerini oldukça dinamik bir araç ortamında optimize etmeyi amaçlamaktadır. Önerilen yapı, karasal baz istasyonları ve yol kenarı birimlerini tamamlayıcı nitelikte olan HAPS ile geniş kapsama alanı ve daha güçlü görüş hattı iletişimi sağlamaktadır.

Kaynak tahsisi, her biri bir araç kümesine karşılık gelen çok sayıda ajanın yer aldığı bir optimizasyon problemi olarak formüle edilmiştir. Ortamın karmaşıklığını yönetmek amacıyla ajanlar, zaman içinde optimal iletişim stratejilerini öğrenmek için MADDPG algoritmasını kullanmaktadır. Ödül fonksiyonu; bilgi yaşı (age of information, AoI) minimize eden, kanal kapasitesini gözetken ve iletim gücü sınırlamalarını dikkate alan çok kriterli bir yapı ile tasarlanmıştır. Bu sayede ajanlar, bilgi tazeliği ile kaynak verimliliği arasında denge kuran politikaları birlikte öğrenebilmektedir.

Simülasyon sonuçları, önerilen HAPS destekli MADDPG çerçevesinin, klasik ve tamamen dağıtık yaklaşımlara kıyasla yakınsama hızı, kararlılık ve genel ödül performansı açısından üstün olduğunu göstermektedir. Özellikle HAPS entegrasyonu sayesinde, değişen araçlar arası mesafelerde daha kararlı ve güncel veri iletimi sağlanarak AoI performansı önemli ölçüde iyileştirilmiştir. Karşılaştırmalı değerlendirmeler, önerilen yöntemin çevresel dalgalanmalara karşı güçlü bir uyarlabilirlik sunduğunu ve sınırlı ağ koşullarında üstün iletişim kalitesi sağladığını ortaya koymaktadır.

Genel olarak, bu çalışma HAPS sistemlerinin ve derin pekiştirmeli öğrenme yaklaşımlarının bütünleşik kullanımının, yeni nesil araç haberleşme ağlarında akıllı ve ölçeklenebilir kaynak yönetimi sağlama potansiyelini güçlü bir şekilde ortaya koymaktadır.

**Anahtar Kelimeler:** Bilgi Yaşı (AoI), Derin Pekiştirmeli Öğrenme (DRL), Yüksek İrtifa Platform İstasyonları (HAPS), Kaynak Yönetimi, Araçtan-Herşeye (V2X) Haberleşme

## ABSTRACT

### MS THESIS

# DEEP LEARNING BASED AND AGE OF INFORMATION AWARE RESOURCE MANAGEMENT IN HAPS-AIDED V2X SYSTEMS

Ahmet Melih İNCE

Konya Technical University  
Institute of Graduate Studies  
Department of Electrical and Electronics Engineering

Advisor: Assoc. Prof. Dr. Ayşe Elif CANBİLEN

2025, 64 Pages

Jury

Advisor Assoc. Prof. Dr. Ayşe Elif CANBİLEN  
Assoc. Prof. Dr. Akif DURDU  
Assoc. Prof. Dr. Hakkı SOY

In this thesis, a novel resource management framework based on the multi-agent deep deterministic policy gradient (MADDPG) algorithm is proposed for high altitude platform station (HAPS)-enabled vehicle-to-everything (V2X) communication systems. The developed framework aims to optimize key performance indicators such as information freshness, energy efficiency, and communication capacity in highly dynamic vehicular environments. The proposed system integrates HAPS into the communication infrastructure as a complementary component to terrestrial base stations and roadside units, providing wider coverage and enhanced line-of-sight connectivity.

Resource allocation problem is formulated as an optimization problem involving multiple agents, each representing a vehicle platoon. To handle the complexity of the environment, agents utilize the MADDPG algorithm to learn optimal communication strategies over time. The reward function is designed as a multi-objective structure that minimizes age-of-information (AoI), satisfies channel capacity constraints, and respects transmission power limits. This design enables agents to jointly learn policies that balance information freshness and resource efficiency.

Simulation results show that the proposed HAPS-assisted MADDPG framework significantly outperforms both classical and fully decentralized approaches in terms of convergence speed, stability, and overall reward performance. In particular, HAPS integration allows for more stable and up-to-date data transmission across varying inter-vehicle distances, leading to significant improvements in AoI performance. Comparative evaluations also demonstrate that the proposed method offers strong adaptability to environmental dynamics and ensures superior communication quality under constrained network conditions.

**Keywords:** Age of Information (AoI), Deep Reinforcement Learning (DRL), High Altitude Platform Station (HAPS), Resource Management, Vehicle-to-Everything (V2X) Communication

## ÖNSÖZ

Tez sürecim boyunca bilgi, tecrübe ve yönlendirmeleriyle katkı sunan danışmanım Doç. Dr. Ayşe Elif CANBİLEN'e en içten teşekkürlerimi sunarım. Gerek akademik rehberliği gerekse her aşamada gösterdiği sabır ve destek, bu çalışmanın ortaya çıkmasında belirleyici rol oynamıştır.

Ayrıca, HAPS sistemleri konusundaki vizyoner fikirleriyle tez çalışmamın gelişimine katkıda bulunan, tez konuyla ilgili Carleton Üniversitesi NTN laboratuvarında gerçekleştirdiğim araştırmalarda desteklerini esirgemeyen Prof. Dr. Halim YANIKÖMEROĞLU'na en derin şükranlarımı sunarım. Bu süreçte Study in Canada Scholarship (SICS) bursu ile araştırma faaliyetlerimin gerçekleşmesine maddi destek ve imkan sağlayan Global Affairs Canada kurumuna da teşekkür ederim.

Beni her zaman destekleyen ve motive eden başta sevgili eşim Merve Melisa İNCE olmak üzere, aileme ve çalışma arkadaşlarıma da içtenlikle teşekkür ederim. Bu tez, onların varlığı sayesinde anlam kazanmıştır.

Bu çalışmanın, HAPS tabanlı haberleşme sistemlerinin geliştirilmesi ve pekiştirmeli öğrenme tekniklerinin mühendislik alanlarındaki uygulamalarına katkı sunması temennisiyle...

Ahmet Melih İNCE  
KONYA-2025

# İÇİNDEKİLER

<b>ÖZET</b> .....	<b>iii</b>
<b>ABSTRACT</b> .....	<b>iv</b>
<b>ÖNSÖZ</b> .....	<b>v</b>
<b>İÇİNDEKİLER</b> .....	<b>vi</b>
<b>SİMGELER VE KISALTMALAR</b> .....	<b>viii</b>
<b>1. GİRİŞ</b> .....	<b>1</b>
<b>2. KAYNAK ARAŞTIRMASI</b> .....	<b>3</b>
2.1. V2X ve HAPS Tabanlı İletişim.....	4
2.2. Kaynak Yönetimi ve DRL Yaklaşımları.....	6
2.2.1. V2X Sistemlerinde Kaynak Yönetimi Problemi.....	7
2.2.2. Kaynak Yönetimi Yöntemleri ve Kısıtları.....	8
2.2.3. AI Tabanlı Uyarlanabilir Yaklaşımlara Geçiş.....	10
2.2.4. DRL ile Yüksek Boyutlu Karar Problemlerinin Çözümü.....	12
2.2.5. Çok Ajanlı Ortamlar İçin MADRL Yaklaşımı.....	14
<b>3. MATERYAL VE YÖNTEM</b> .....	<b>16</b>
3.1. V2X İletişim Teknolojileri ve Mimarileri.....	16
3.2. AoI Kavramı.....	19
3.3. HAPS Sistemleri ve V2X Entegrasyonu.....	21
3.4. V2X'te Kaynak Yönetimi Yöntemleri.....	24
3.4.1. Geleneksel Kaynak Yönetimi Yöntemleri ve Sınırları.....	25
3.4.2. Güncel Kaynak Yönetimi Yöntemleri.....	27
3.5. Derin Takviyeli Öğrenme (DRL).....	29
3.5.1. MADRL ve MADDPG Yaklaşımı.....	30
<b>4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA</b> .....	<b>33</b>
4.1. Sistem Mimarisi.....	33
4.2. Kanal Modeli.....	35
4.2.1. Karasal Bağlantı Kanalları (V2V ve V2I).....	35
4.2.2. HAPS Bağlantısı (V2H).....	36
4.2.3. Bağlantı Eşiği ve Kanal Kullanılabilirliği.....	38
4.3. AoI Modeli.....	39
4.3.1. İletişim Kapasitesi.....	42
4.3.2. AoI Güncelleme Kuralları.....	45
4.3.3. Optimizasyon Probleminin Formülizasyonu.....	46
4.4. MADDPG Tabanlı Çözüm.....	47
<b>5. SONUÇLAR VE ÖNERİLER</b> .....	<b>52</b>

5.1. Sonular.....	52
5.2. neriler.....	59
<b>KAYNAKLAR.....</b>	<b>61</b>



## SİMGELER VE KISALTMALAR

### Simgeler

$\Delta_i(t)$	: $i$ numaralı ajan için $t$ anındaki AoI değeri
$\bar{\Delta}$	: Ortalama AoI
$\Delta_{max}$	: Maksimum AoI değeri
$P_t$	: $t$ anındaki iletim gücü
$h_k$	: $k$ kanalının kazancı
$\zeta_j$	: $j$ ajanı için kapasite eşiği
$\delta$	: Kronecker delta fonksiyonu
$\pi_i$	: $i$ ajanına ait politika
$\theta_j$	: $j$ ajanının seçtiği iletişim modu
$r$	: Ödül fonksiyonu
$\mathcal{A}$	: Eylem uzayı
$\mathcal{S}$	: Durum uzayı
$\kappa_1, \kappa_2, \dots$	: Ödül fonksiyonu ağırlık katsayıları
$K_1, K_2, K_3, K_4$	: Optimizasyon kısıtları
$J(\pi_i)$	: Politika performansı (beklenen ödül)
$\mathbb{E}$	: Beklenen değer operatörü
$\mathbb{P}$	: Olasılık operatörü

### Kısaltmalar

AI	: Yapay Zeka (Artificial Intelligence)
AoI	: Bilgi Yaşı (Age of Information)
C-V2X	: Araçtan Her Şeye Hücresel Bağlantı (Cellular Vehicle-to-Everything)
dB	: Desibel
DDPG	: Derin Deterministik Politika Gradyanı (Deep Deterministic Policy Gradient)
DNN	: Derin Sinir Ağları (Deep Neural Network)

DQN	: Derin Q Ağları (Deep Q-Network)
DRL	: Derin Pekiştirmeli Öğrenme (Deep Reinforcement Learning)
DSRC	: Özel Kısa Menzilli İletişimler (Dedicated Short-range Communications)
HAPS	: Yüksek İrtifa Platform İstasyonu (High Altitude Platform Station)
HRLLC	: Hiper Güvenilir Düşük Gecikmeli İletişim (Hyper-Reliable Low Latency Communication)
ITS	: Akıllı Ulaşım Sistemleri (Intelligent Transportation Systems)
LoS	: Doğrudan Görüş (Line of Sight)
MADDPG	: Çoklu Ajan Derin Deterministik Politika Gradyanı (Multi-Agent Deep Deterministic Policy Gradient)
MADRL	: Çoklu Ajan Derin Takviyeli Öğrenme (Multi-Agent Deep Reinforcement Learning)
ML	: Makine Öğrenmesi (Machine Learning)
LTE	: Uzun Vadeli Evrim (Long-Term Evolution)
NTN	: Karasal Olmayan Ağ (Non-Terrestrial Network)
NR-V2X	: Araçtan Herşeye Yeni Radyo (New Radio Vehicle-to-Everything)
OFDM	: Ortogonal Frekans Bölmeli Çoklama (Orthogonal frequency-division multiplexing)
PL	: Filo Lideri (Platoon Leader)
RL	: Pekiştirmeli Öğrenme (Reinforcement Learning)
RSU	: Yol Kenarı Birimi (Road Side Unit)
SINR	: Sinyal girişim gürültü oranı (Signal-to-Interference-plus-Noise Ratio)
QoS	: Hizmet Kalitesi (Quality of Service)
V2H	: Araçtan-HAPS'e (Vehicle-to-HAPS)
V2I	: Araçtan-Altyapıya (Vehicle-to-Infrastructure)
V2N	: Araçtan-Sunuculara (Vehicle-to-Network)
V2V	: Araçtan-Araca (Vehicle-to-Vehicle)
V2X	: Araçtan-Her şeye (Vehicle-to-Everything)
3GPP	: Üçüncü Nesil Ortaklık Projesi (Third Generation Partnership Project)
6G	: Altıncı Nesil (Sixth Generation)



## 1. GİRİŞ

Otonom sistemlerin hızla geliştiği günümüzde, ulaşım altyapıları da bu dönüşüm sürecine paralel olarak yeniden şekillenmektedir. Bu dönüşüm, yalnızca sürücüsüz araç teknolojileri gibi uç uygulamaları değil; aynı zamanda trafik güvenliği, bilgi güncelliği, enerji verimliliği ve iletişim sürekliliği gibi temel sistem performanslarını da doğrudan etkilemektedir. Bu çerçevede geliştirilen araçtan-her şeye (vehicle-to-everything, V2X) iletişim sistemleri, hem bireysel araçlar arasında hem de altyapı ve bulut birimleriyle gerçek zamanlı veri alışverişi kurarak yeni nesil ulaşım ağlarının temelini oluşturmaktadır.

V2X sistemlerinin başarısı, haberleşme ağlarının sunduğu gecikme, bant genişliği ve güvenilirlik gibi metriklerle birlikte, son yıllarda önem kazanan güncellik odaklı metriklere de bağlıdır. Bilgi Yaşı (age of information, AoI), alıcı tarafından kullanılan bilginin ne kadar “taze” olduğunu ifade eder ve özellikle yüksek hareketlilik içeren V2X ortamlarında sistem kararlarının doğruluğunu belirleyen kritik bir göstergedir. Ancak bu tür zaman duyarlı bir bilgi akışının sağlanabilmesi için yalnızca protokol düzeyinde çözümler yeterli olmamakta; aynı zamanda ağın esnek, ölçeklenebilir ve bağlamsal koşullara adapte olabilen bir kaynak yönetimi yapısına da sahip olması gerekmektedir.

Bu bağlamda geleneksel karasal haberleşme altyapıları yetersiz kalabilmekte ve bu da iletişim sürekliliğini tehdit etmektedir. Bu tür boşlukların doldurulabilmesi amacıyla geliştirilen karasal olmayan ağ (non-terrestrial network, NTN) çözümleri arasında, stratosfere yerleştirilen yüksek irtifa platform istasyonları (high altitude platform station, HAPS) öne çıkmaktadır. HAPS geniş kapsama alanı, düşük gecikme ve daha güçlü görüş hattı bağlantısı sunarak V2X sistemlerine stratejik bir destek bileşeni olarak entegre edilebilmektedir.

Ancak HAPS tabanlı V2X mimarisinin sunduğu bu avantajlardan etkin bir şekilde yararlanılabilmesi için sistemin kaynak yönetimi düzeyinde akıllı ve bağlama duyarlı kararlar alabilmesi gerekmektedir. Bu noktada, klasik sabit eşiklere dayalı merkezi kontrol sistemleri yetersiz kalmakta ve yerini öğrenebilen, dağıtık ve otonom yapıdaki yapay zeka (artificial intelligence, AI) çözümlerine bırakılmaktadır. Özellikle derin pekiştirmeli öğrenme (deep reinforcement learning, DRL) tabanlı algoritmalar, bu

tür ortamlarda karar verici ajanların çevreye adapte olabilmelerini ve uzun vadeli sistem optimizasyonunu mümkün kılmaktadır.

Bu tez çalışması, HAPS destekli V2X haberleşme sistemlerinde bilgi güncelliğini merkeze alan, dinamik ve çok ajanlı bir kaynak yönetimi yaklaşımı geliştirmeyi amaçlamaktadır. Bu kapsamda, sürekli aksiyon uzayları ve çok boyutlu karar süreçleri için uygunluğu kanıtlanmış olan çoklu ajan derin deterministik politika gradyanı (multi-agent deep deterministic policy gradient, MADDPG algoritması esas alınmıştır. Her bir araç, bağımsız bir “öğrenen ajan” olarak modellenmiş olup; bu ajanlar hem kendi gözlemlerine dayalı olarak aksiyon üretmekte hem de merkezi bir eleştirilen tarafından değerlendirilmektedir. Böylece dağıtık karar alma ile merkezi değerlendirme arasında bir denge kurularak, sistem genelinde bilgi tazeliği, iletişim kalitesi ve enerji verimliliği birlikte optimize edilmeye çalışılmıştır.

Bu çerçevede, HAPS tabanlı V2X sistemleri için geliştirilen MADDPG destekli öğrenme çerçevesi, klasik ve dağıtık yaklaşımlarla karşılaştırmalı olarak analiz edilmiş; simülasyon sonuçları üzerinden yöntem performansı değerlendirilmiştir. Elde edilen bulgular, önerilen yaklaşımın, özellikle değişken kanal koşulları ve yoğun araç trafiği senaryolarında üstün performans sergilediğini ve yeni nesil otonom iletişim sistemleri için güçlü bir alternatif sunduğunu ortaya koymaktadır.

## 2. KAYNAK ARAŞTIRMASI

Dijital dönüşümün etkisiyle birlikte, ulaşım sektörü ciddi bir evrim sürecinden geçmektedir. Bu dönüşüm, yalnızca sürücülerin deneyimini dönüştürmekle kalmayıp, aynı zamanda trafik güvenliği, trafik akışının optimizasyonu ve enerji verimliliği gibi pek çok önemli alanda da yeni olanaklar yaratmaktadır. Bu çerçevede geliştirilen akıllı ulaşım sistemleri (intelligent transportation systems, ITS) ve özellikle V2X teknolojileri, gelecek nesil ulaşım altyapılarının temel yapı taşlarını oluşturmaktadır (Clancy ve ark., 2024).

V2X kavramı, yalnızca araçtan-araca (vehicle-to-vehicle, V2V) değil, aynı zamanda araçtan-altyapıya (vehicle-to-infrastructure, V2I), yayalarla (vehicle-to-pedestrian, V2P) ve merkezi sunucularla (vehicle-to-network, V2N) sürekli bilgi alışverişinde bulunduğu bir iletişim modelini ifade eder. Bu modelin amacı, araçların çevresel farkındalık düzeyini artırarak, olası tehlikelerin önceden tespit edilmesini ve daha güvenli bir sürüş deneyimi sağlanmasını mümkün kılmaktır (Hasan ve ark., 2024).

Özellikle altıncı nesil (sixth generation, 6G) kablosuz iletişim teknolojilerinin gündeme gelmesiyle birlikte, V2X sistemlerinden beklenen performans kriterleri de yeniden tanımlanmıştır. 6G'nin sunduğu hiper güvenilir düşük gecikmeli iletişim (hyper-reliable low latency communication, HRLLC) vizyonu; milisaniyeler düzeyinde tepki sürelerini, yüksek veri aktarım oranlarını ve sürekli bağlantıyı gerektirir. Bu durum, özellikle otonom sürüş sistemleri, çarpışma önleyici mekanizmalar ve gerçek zamanlı trafik koordinasyonu gibi güvenlik-kritik uygulamalarda hayati bir öneme sahiptir.

Bu tür sistemlerin başarısı büyük ölçüde veri güncelliğine, yani sistemin karar alırken ne kadar güncel bilgi kullandığına bağlıdır. Bu noktada, iletişim ağlarının performansını ölçmek için geleneksel metrikler olan gecikme (latency) ve bant genişliği gibi ölçütlerin yanında, son yıllarda literatürde önem kazanan bir başka metrik olarak AoI öne çıkmaktadır. AoI, bir bilginin gönderildiği andan, alıcı tarafından kullanıldığı ana kadar geçen süreyi temsil eder ve özellikle dinamik sistemlerde kararların doğruluğunu ve güncelliğini doğrudan etkileyen bir ölçüt olarak kabul edilmektedir (Wang ve ark., 2024). Ancak, bu tür bilgi güncelliğinin sağlanabilmesi için yalnızca iletişim protokollerinin sağlanması yeterli olmaz. Aynı zamanda, bu bilgilerin sürekli ve kesintisiz bir iletişim altyapısı üzerinden aktarılabilmesi gerekmektedir. Burada, karasal

ağ altyapıları çoğu zaman yetersiz kalmaktadır. Özellikle altyapı yoğunluğunun düşük olduğu kırsal bölgeler, doğal afetler sonrası bağlantı kesintisi yaşanan alanlar veya hızlı hareket eden araç kümelerinin yoğun olduğu otoyol senaryoları, bu zorlukların başında gelmektedir.

Bu tür iletişim boşluklarını doldurmak amacıyla geliştirilen NTN çözümleri arasında HAPS öne çıkmaktadır. HAPS sistemleri, stratosfer katmanında konumlanan, güneş enerjisiyle çalışan ve uzun süre havada kalabilen platformlar olup, hem geniş kapsama alanı hem de karasal altyapıdan bağımsız çalışma esnekliği ile V2X sistemlerine yeni bir boyut kazandırmaktadır. HAPS, geleneksel baz istasyonlarına kıyasla daha yüksek doğrudan görüş (line of sight, LoS) olasılığı sunduğu için, bilgi güncelliği açısından büyük avantajlar sağlamaktadır (Abbasi ve ark., 2024).

Bu bağlamda, bu tez çalışması; 6G hedefleri doğrultusunda, V2X sistemlerinin bilgi güncelliğini artırmak ve iletişim sürekliliğini sağlamak amacıyla HAPS teknolojisinin nasıl entegre edilebileceğini ve bu sürecin nasıl optimize edilebileceğini araştırmaktadır. Böylece hem iletişim altyapısının esnekliği artırılmakta, hem de araçların çevresel farkındalık düzeyi güçlendirilerek otonom sürüş sistemlerinin güvenilirliği artırılmaktadır.

## 2.1. V2X ve HAPS Tabanlı İletişim

Haberleşme sistemlerinin kapsamı ve önemi her geçen gün artmaktadır. Özellikle gerçek zamanlı, kesintisiz ve güvenilir veri iletiminin kritik hâle geldiği ulaşım uygulamaları, bu ihtiyacın en net biçimde hissedildiği alanların başında gelmektedir. V2X sistemleri, bu kapsamda yalnızca bireysel veri aktarımına dayalı bir ağ olmanın ötesine geçerek, çok katmanlı ve çok kaynaklı bir iletişim mimarisine dönüşmektedir. Ancak bu dönüşüm, karasal haberleşme altyapılarıyla sınırlandırıldığında, özellikle kapsama alanı ve erişim sürekliliği açısından ciddi darboğazlara neden olmaktadır.

Karasal ağ mimarileri; baz istasyonları, yol kenarı birimi (road side unit, RSU) ve yerel sunucular aracılığıyla V2X sistemlerini desteklemeye çalışmaktadır. Ancak bu yapılar, fiziki altyapı eksikliği, yüksek kurulum maliyeti, bakım zorlukları ve doğal afet riski gibi nedenlerle özellikle kırsal bölgelerde, dağlık alanlarda veya acil müdahale gereken durumlarda sınırlı bir hizmet sunabilmektedir (Kurt ve ark., 2021). Bu

bağlamda, V2X iletişim sistemlerinin gerektirdiği düşük gecikmeli ve yüksek güvenilirlikli veri alışverişi, bu tür zayıf kapsamalarda sağlıklı bir şekilde sürdürülememektedir. Özellikle çarpışma önleyici sistemler, otonom sürüş kararları ve trafik akış koordinasyonu gibi uygulamalar için iletişimin anlık kesintiye uğraması bile sistemin güvenliğini doğrudan riske atmaktadır.

Bu tür altyapı zayıflıklarını telafi etmek amacıyla geliştirilen NTN yapıları, yeryüzü ile gökyüzü arasında farklı yükseklik katmanlarında haberleşme platformlarının konumlandırılması yoluyla, daha kapsayıcı ve esnek bir iletişim mimarisi sunmayı hedeflemektedir (Giordani ve ark., 2020). HAPS, hem karasal hem de uydu tabanlı sistemlerin avantajlarını birleştiren hibrit bir çözüm sunarak, V2X gibi dinamik ve yaygın sistemlerin ihtiyaç duyduğu geniş kapsama, düşük gecikme ve yüksek bağlantı kalitesini aynı anda karşılayabilme kapasitesine sahiptir (Kurt ve ark., 2021).

HAPS sistemlerinin en temel avantajlarından biri, LoS sağlayabilen yüksek irtifa konumları sayesinde, daha az sinyal kaybıyla daha geniş alanlara hizmet sunabilmesidir. Bu özellik, özellikle yoğun şehir yapılaşması veya dağlık coğrafyalarda sinyal tıkanıklığını ve çoklu yansıma etkilerini minimize ederek, veri iletiminin daha kararlı ve güçlü olmasını sağlamaktadır. Tipik bir HAPS platformu, yüzlerce kilometrekarelik bir alanı kapsayabilme yeteneğine sahip olup; bu sayede hem kırsal alanlarda altyapı ihtiyacını ortadan kaldırmakta hem de acil durum senaryolarında hızlı ve geçici ağ altyapıları oluşturulmasına imkân vermektedir.

Buna ek olarak, HAPS sistemleri 6G iletişim vizyonu doğrultusunda şekillenen geleceğin ağ altyapılarına da uyumlu niteliktedir. 6G'nin temel hedeflerinden biri olan HRLLC, mevcut karasal altyapılarla sınırlı kalındığında pek çok senaryoda karşılanamamaktadır. Ancak HAPS, gerek düşük yörüngede bulunması nedeniyle, uydulara kıyasla sağladığı düşük gecikme süresi, gerekse insansız hava araçlarına kıyasla sağladığı geniş kapsama alanı sayesinde bu gereksinimleri karşılayabilecek nadir çözümlerden biridir. Bu kapsamda HAPS, karasal baz istasyonları ile uydu sistemleri arasında konumlanan bir ara bağlantı katmanı olarak değerlendirilmekte ve ağ mimarisine katmanlı bir yapı kazandırarak esneklik sağlamaktadır.

V2X sistemlerine özel olarak bakıldığında, HAPS'in en dikkat çekici katkılarından biri, ağ yedekliliği ve kesintisiz bağlantı sağlamasıdır (Lou ve ark., 2023). Karasal baz istasyonlarının devre dışı kaldığı senaryolarda, HAPS devreye girerek veri iletimini sürdürebilir ve özellikle güvenlik-kritik haberleşme (örneğin araçlar arası

çarpışma önleyici uyarılar) için yedek kanal oluşturabilir. Bu sayede, otonom araçlar ve bağlantılı taşıma sistemleri, sadece altyapıya bağımlı kalmadan karar alma süreçlerini sürdürebilir hale gelmektedir.

Bununla birlikte, HAPS sistemleri yalnızca fiziksel erişim sunmakla kalmayıp, aynı zamanda spektrum verimliliği ve dinamik kaynak yönetimi açısından da avantaj sağlamaktadır. Yönlendirilebilir anten yapıları ve bölgesel hücre oluşturma kapasiteleri sayesinde, farklı bölgelerdeki trafik yoğunluğu göz önünde bulundurularak spektrum kaynakları talebe göre yeniden tahsis edilebilir. Bu dinamik yapı, spektrum paylaşımının daha verimli gerçekleşmesini sağlayarak, yoğun trafiğe sahip bölgelerde bile gecikme ve çakışma sorunlarını minimuma indirebilir.

HAPS tabanlı iletişim mimarileri, yalnızca birer yedekleme çözümü değil; aynı zamanda V2X sistemlerinin güvenliği, esnekliği ve sürekliliği açısından stratejik birer bileşen olarak değerlendirilmektedir. 6G sonrası iletişim paradigmasında HAPS'ın entegrasyonu, sadece kapsama alanını artırmakla kalmayacak, aynı zamanda ağın zekâsını, adaptasyon kabiliyetini ve veri güncelliğini de önemli ölçüde geliştirecektir (Abbasi ve ark., 2024). Bu nedenle, V2X sistemlerinin geleceği HAPS gibi yeni nesil NTN çözümleriyle birlikte değerlendirilmeli ve bu çözümler doğrultusunda yeniden yapılandırılmalıdır.

## 2.2. Kaynak Yönetimi ve DRL Yaklaşımları

Günümüzün yoğun ve dinamik V2X haberleşme ortamlarında, sınırlı ağ kaynaklarının verimli, adil ve bağlama duyarlı biçimde yönetilmesi, sistem performansı açısından kritik bir gerekliliktir. Bu bölümde, V2X sistemlerinde kaynak yönetimi problemi detaylı biçimde incelenmekte; geleneksel yöntemlerin sınırlılıkları ortaya konularak, AI tabanlı uyarlanabilir yaklaşımlara geçiş ihtiyacı temellendirilmektedir. Özellikle pekiştirmeli öğrenme (reinforcement learning, RL) gibi modern AI tekniklerinin, V2X sistemlerine sağladığı esneklik ve bağlamsal farkındalık kabiliyeti üzerinde durulmaktadır. Ayrıca çok ajanlı ortamlar için geliştirilen yaklaşımlar değerlendirilerek, bu yapılarla yüksek boyutlu, sürekli ve etkileşimli karar problemlerinin nasıl yönetilebileceği tartışılmaktadır.

### 2.2.1. V2X Sistemlerinde Kaynak Yönetimi Problemi

V2X, modern ulaşım altyapılarında gerçek zamanlı veri paylaşımını mümkün kılan temel bir teknolojidir. Bu iletişim yapısı, yalnızca V2V değil, aynı zamanda V2I, V2P, V2N ve hatta V2H ile sürekli bilgi alışverişinde bulunmasına imkân tanıyabilir. V2X sistemleri, otonom sürüş teknolojileri, çarpışma önleme sistemleri, trafik yoğunluğu yönetimi ve yol güvenliği gibi çok sayıda kritik uygulamanın çalışmasını doğrudan etkiler. Bu nedenle, sistemin iletişim sürekliliği, veri güncelliği, düşük gecikme ve yüksek güvenilirlik gibi performans kriterlerini karşılaması büyük önem taşır. Bu performans kriterlerinin sağlanabilmesi için, sistemde mevcut olan ağ kaynaklarının etkin ve dengeli biçimde yönetilmesi gerekir. Burada söz konusu olan kaynaklar, başta spektrum (frekans bandı) olmak üzere; iletişim zamanı, iletim gücü, alt kanal erişimi, veri oranları ve ağ kapasitesi gibi unsurları kapsar. Bu kaynakların tamamı sınırlı yapıya sahip olduğu için, çok sayıda aracın eşzamanlı iletişim kurmaya çalıştığı ortamlarda bu kaynaklar arasında çatışma ya da aşırı yüklenme durumu kaçınılmaz hâle gelir (Bazzi ve ark., 2021).

V2X sistemleri, diğer haberleşme sistemlerinden farklı olarak, yüksek hareketlilik ve dinamik topoloji değişimleri ile karakterize edilir. Özellikle araçların farklı hızlarda hareket etmesi, anlık manevralar yapması, ağdan çıkıp tekrar dahil olması gibi durumlar, iletişim ortamının deterministik olmayan bir yapıya sahip olmasına neden olur. Bu koşullarda, ağ kaynaklarının etkin tahsisi oldukça karmaşık bir optimizasyon problemine dönüşmektedir. Örneğin bir aracın aniden yavaşlaması veya kavşağa yaklaşması, yalnızca kendi iletişim ihtiyacını değil; çevresindeki tüm araçlarla olan iletişim akışını da etkileyebilir. Dolayısıyla, kaynak yönetimi yalnızca bireysel değil, kooperatif karar alma süreçlerini de içermek zorundadır.

Bir diğer önemli unsur da heterojen iletişim modlarının bir arada bulunmasıdır. V2X ortamlarında V2V, V2I, V2H gibi farklı haberleşme yolları aynı anda devrede olabilir. Bu durum, her bir aracın anlık bağlamına ve ortam koşullarına bağlı olarak hangi mod üzerinden iletişim kurması gerektiğine karar verilmesini zorlaştırır. Örneğin şehir merkezinde RSU ile bağlantı sağlanabilirken, kırsal bölgede bu görev HAPS tarafından üstlenilebilir. Bu bağlamda kaynak yönetimi, yalnızca fiziksel kaynakların paylaşımını değil; aynı zamanda iletişim yolunun seçimi gibi karmaşık kararları da kapsar.

Özellikle çoklu araç filosu (platoon) yapılarında, kaynak paylaşımı daha da kritik hâle gelir. Araçlar, küme lideri ve takipçiler arasında sürekli bilgi paylaşımı gerçekleştirirken, diğer kümelerle olan iletişimini de sürdürmek zorundadır. Bu da hem kendi içindeki (intra-platoon) hem de diğer kümelerle (inter-platoon) iletişimi optimize etmeyi gerektirir. Yani bir aracın hangi kaynaklara ne zaman ve hangi iletişim modu üzerinden erişeceği; yalnızca kendi performansı değil, tüm sistemin verimliliği açısından belirleyici olur (Parvini ve ark., 2023).

Bu bağlamda, V2X sistemlerinde kaynak yönetimi problemi; çok boyutlu, dinamik, dağıtık ve zaman duyarlı bir karar alma sürecine dönüşmektedir. Söz konusu kaynakların optimum şekilde kullanılması, sadece ağın performansını artırmakla kalmaz; aynı zamanda AoI, enerji verimliliği ve güvenlik gibi üst düzey hedeflere ulaşmak için de gereklidir.

Bu sebeple, kaynak yönetimi problemi V2X sistemlerinin temel zorluklarından biri hâline gelmiş olup, hem endüstriyel uygulamalar hem de akademik çalışmalar açısından çözülmesi gereken öncelikli konular arasında yer almaktadır.

### **2.2.2. Kaynak Yönetimi Yöntemleri ve Kısıtları**

V2X sistemlerinde kaynak yönetimi problemi, ilk dönemlerde çoğunlukla statik, merkezi ve kurala dayalı yöntemlerle ele alınmıştır. Bu yöntemler, düşük mobiliteli sistemler ve sabit topolojiler için kabul edilebilir performans gösterse de günümüzdeki yüksek mobiliteli, dinamik, dağıtık ve zaman duyarlı V2X yapıları için ciddi ölçüde yetersiz kalmaktadır. Kaynak tahsisi; iletişim kanalları, zaman dilimleri, spektrum kullanımı ve iletim gücü gibi sınırlı kapasiteye sahip öğelerin, yoğun veri trafiği altında adil ve verimli biçimde paylaşılmasını gerektirir. Bu ihtiyacı karşılamak amacıyla geleneksel sistemlerde genellikle sabit eşik değerlerine, önceden belirlenmiş senaryolara veya merkezi kontrol birimlerine dayalı tahsis algoritmaları kullanılmaktadır (Sehla ve ark., 2022).

Sabit kaynak tahsisi yaklaşımlarında, her araca ya da araç grubuna önceden tanımlanmış kanal, zaman dilimi veya frekans bandı atanır. Bu yöntemler uygulama açısından basittir, ancak değişken trafik yoğunluklarına veya anlık kanal kalitesine karşı duyarsız olduklarından dolayı kaynakların ya verimsiz kullanılmasına ya da bazı kullanıcıların kaynaklara hiç erişememesine yol açar. Yoğun trafik zamanlarında oluşan

aşırı yüklenme durumu gecikmelerin artmasına, paket kayıplarına ve bilgi güncelliğinin azalmasına neden olurken; düşük yoğunluk zamanlarında sistemin kapasitesi atıl kalır. Bu durum, özellikle bilgi güncelliğinin kritik olduğu güvenlik uygulamaları için kabul edilemez bir performans kaybı anlamına gelir (Parvini ve ark., 2023).

Geleneksel sistemlerde sıklıkla kullanılan bir diğer yöntem, öncelik tabanlı kaynak tahsisidir. Bu yöntemde, iletişim paketleri önem derecelerine göre sınıflandırılır ve kaynaklar önceden belirlenmiş öncelik sırasına göre dağıtılır. Örneğin acil frenleme uyarıları veya çarpışma riski bildiren mesajlara daha yüksek öncelik verilerek hızlı iletim sağlanmaya çalışılır. Ancak bu tür sabit öncelik yapıları, düşük öncelikli mesajların sistematik olarak göz ardı edilmesine neden olabilir. Trafik akışını destekleyen, konforla ilgili veya enerji verimliliği sağlayan bilgi paketleri zamanla sistemden dışlanır ve bu durum, uzun vadede sistemin bütünsel verimliliğini olumsuz etkiler (Li ve ark., 2022).

Merkezi kontrol sistemleri ise tüm ağı izleyebilen birimlerin (örneğin baz istasyonu, RSU veya bulut sunucusu) karar verici olduğu yapılardır. Bu sistemler, teoride daha bütüncül bir değerlendirme yapma imkânına sahip oldukları için kaynakları daha global bir bakış açısıyla tahsis edebilirler. Ancak yüksek mobilitelere sahip V2X ortamlarında, bu yapıların da ciddi zaafı vardır (Vu ve ark., 2022). Öncelikle ağın her bir bileşeninin merkezle sürekli olarak iletişim hâlinde olması gerektiğinden, bu yapı yüksek gecikme üretir. Özellikle otonom sürüş uygulamaları gibi milisaniyelik kararların kritik olduğu senaryolarda, merkezi sistemlerin tepki süresi, güvenlik açısından kabul edilemez seviyelere ulaşabilir. Ayrıca merkezi mimariler, sistemde bir arıza noktası oluştuğunda tüm iletişimin aksamasına neden olabilir. Bu durum, V2X sistemlerinin yaygınlaştırılmasında güvenlik ve süreklilik açısından ciddi bir engeldir. Merkezi kontrolün ölçeklenebilirliği de ciddi bir başka sorundur. Araç sayısının arttığı ortamlarda merkez birimin yükü artmakta, bu da işlem süresini ve karar gecikmesini artırarak performansı olumsuz etkilemektedir. Sistem kontrol yükünün merkezi bir yapıya aşırı bağlı olması, ağ trafiğinde tıkanıklıklara ve iletimde tutarsızlıklara yol açmaktadır.

Buna ek olarak, geleneksel yöntemler kanal kalitesindeki ani değişimlere uyum sağlayamaz. Sabit kanal tahsisi yapılarına sahip sistemlerde, kanal ortamının bozulması durumunda araçlar hâlâ düşük kaliteli bağlantıları kullanmaya devam edebilir. Bu da iletim başarımını düşürür ve özellikle zaman kritik veri paketlerinin kaybolmasına

neden olabilir. Aynı şekilde, sabit iletim gücü kullanan sistemler ya gereksiz yere fazla enerji tüketir ya da iletişim başarısızlıklarıyla karşılaşır. Oysa V2X sistemlerinde kanal kazancı, ortam koşulları, araçların konumları, hızları ve yönleri gibi faktörlere bağlı olarak hızla değişebilmektedir (Yoshizawa ve ark., 2023). Geleneksel sabit eşikli sistemler, bu değişkenliğe duyarsızdır ve çevre koşullarını analiz etmeden karar verirler.

Öte yandan, V2X sistemleri yalnızca tek bir iletişim yoluna değil; V2V, V2I, V2H ve V2N gibi birçok haberleşme moduna aynı anda sahip olmalıdır. Bu iletişim yolları ortam koşullarına göre avantaj veya dezavantaj sağlayabilir. Ancak geleneksel yöntemler bu modlar arasında esnek geçiş yapma becerisine sahip değildir. Oysa bazı durumlarda, örneğin karasal ağın yetersiz kaldığı bir bölgede HAPS aracılığıyla haberleşme sağlanması gerekebilir (Kurt ve ark., 2021). Hangi modun ne zaman kullanılacağına dair kararlar geleneksel sistemlerde ya sabit kurallara dayanır ya da hiç yapılmaz. Bu da hem performans kaybına hem de ağ güvenilirliğinin düşmesine neden olur. Geleneksel kaynak yönetimi yaklaşımları, günümüzdeki V2X sistemlerinin ihtiyaç duyduğu esneklik, tepki hızı, bilgi güncelliği, güvenilirlik ve enerji verimliliği gibi temel performans hedeflerini karşılamakta yetersiz kalmaktadır. Hızla değişen iletişim ortamlarına adapte olamayan, sabit kurallara bağlı, çevresel farkındalığı düşük ve karar süreçleri gecikmeli olan bu yöntemlerin yerine, daha esnek, öğrenebilen ve dağıtık yapılar üzerine kurulu yeni nesil karar sistemlerine ihtiyaç duyulmaktadır. Bu bağlamda, AI temelli yaklaşımlar, V2X sistemleri için dinamik, otonom ve bağlama duyarlı kaynak yönetimi mekanizmaları sunma potansiyeline sahiptir.

### **2.2.3. AI Tabanlı Uyarlanabilir Yaklaşımlara Geçiş**

AI tabanlı karar sistemleri, V2X iletişim gereksinimlerini karşılayabilecek potansiyele sahip teknolojik çözümler sunmaktadır. Özellikle makine öğrenmesi (machine learning, ML) ve alt dalı olan RL, V2X ortamlarında karşılaşılan belirsizlikler ve değişkenliklerle baş edebilmek için son derece elverişli araçlardır. RL, ajan adı verilen karar verici birimlerin, çevreyle etkileşime girerek en uygun eylem dizisini öğrenmesini sağlar. Bu öğrenme süreci, ödül (reward) sinyaliyle yönlendirilir ve ajanlar uzun vadeli ödülü maksimize edecek politika fonksiyonlarını geliştirir. En önemli avantajlarından biri, sistemin çevreyi tam olarak bilmesine gerek olmadan, deneme-yanılma ile öğrenebilmesidir (Lai ve ark., 2024).

V2X sistemleri özelinde düşünülürken, her araç veya iletişim birimi birer “öğrenen ajan” olarak modellenilebilir. Bu ajanlar; kanal durumu, trafik yoğunluğu, AoI, iletim başarımı, enerji seviyesi gibi çevresel bilgileri gözlemler. Bu gözlemler doğrultusunda, hangi iletişim yolunun kullanılacağı, hangi kanalda iletim yapılacağı, ne kadar güç kullanılacağı gibi kararları verir. Bu kararlar doğrudan ağaç yapılarına veya sabit eşiklere değil, öğrenilen politika fonksiyonlarına dayanır. Böylece RL tabanlı sistemler, çevresel koşullara göre davranışlarını uyarlayabilen, bağlama duyarlı yapılar hâline gelir (Zhang ve ark., 2019).

Örneğin bir araç, çevresindeki kanal ortamının gürültülü olduğunu gözlemlediğinde iletim gücünü artırabilir ya da karasal bağlantı kopmak üzereyse HAPS destekli V2H kanalına geçiş yapmayı tercih edebilir. Bu tür kararlar geleneksel algoritmalarla ancak önceden programlanmışsa verilebilirken, RL sistemleri bunu öğrenilmiş davranış biçimi olarak kendi kendine gerçekleştirebilir (Liu ve ark., 2020). Bu özelliği sayesinde RL, dinamik ve tahmin edilemeyen koşullarda karar verme konusunda önemli bir avantaj sağlar. RL sistemleri, zaman içinde daha iyi kararlar almak üzere kendilerini geliştirme yeteneğine sahiptir (Arulkumaran ve ark., 2017). Bu durum, ağın uzun vadeli performansını artırmak açısından kritiktir. Klasik sistemlerde performans ancak insan müdahalesiyle güncellenirken, RL sistemleri ağ üzerinde çalışırken aynı zamanda öğrenmeye devam eder. Böylece hem ağ yönetim yükü azalır hem de ağın genel zekâ seviyesi artar.

Bununla birlikte, RL yöntemleri ayrık ve düşük boyutlu aksiyon uzaylarında etkili çalışırken; V2X gibi sürekli değerler içeren, çok boyutlu ve yüksek karmaşıklıkta ortamlarda performans düşüklüğü yaşanabilir (Parvini ve ark., 2023). Örneğin iletim gücünün, kanal bant genişliğinin veya zaman dilimi seçiminin sürekli değerler aldığı bir ortamda, klasik RL teknikleri yetersiz kalabilir. Bu durumda devreye DRL girmektedir. DRL, derin sinir ağlarını kullanarak hem durum uzayının hem de aksiyon uzayının karmaşıklığını yönetebilir. Böylece daha büyük, belirsiz ve gürültülü ortamlar için kararlar üretilebilir.

DRL sistemleri, V2X gibi karmaşık senaryolarda; spektrum tahsisi, iletim gücü optimizasyonu, kanal geçişi, mod seçim stratejileri ve bilgi güncelliğinin korunması gibi görevlerde etkili şekilde kullanılabilir. Ayrıca RL/DRL sistemlerinin dağıtık olarak çalışabilme yetenekleri, merkezi kontrol ihtiyacını azaltır ve tekil hata noktalarını ortadan kaldırır. Araçlar birbirinden bağımsız olarak karar alabilirken, yerel bilgilerle

küresel performansa katkı sağlar. AI tabanlı uyarlanabilir karar sistemleri, V2X ağlarının ihtiyaç duyduğu dinamik kaynak yönetimini gerçekleştirmek için son derece uygun bir zemine sahiptir. Bu yaklaşımlar, sadece algoritmik bir iyileştirme değil, aynı zamanda sistem tasarımı açısından da paradigma değişimini temsil eder. Geleneksel sabit kurallı sistemlerin yerini, bağlamsal farkındalığa sahip, öğrenen ve evrilen sistemlerin alması; hem bilgi güncelliği hem de ağı uzun vadeli sürdürülebilirliği açısından stratejik bir gerekliliktir (Ye ve ark., 2019).

#### **2.2.4. DRL ile Yüksek Boyutlu Karar Problemlerinin Çözümü**

RL algoritmaları, V2X gibi zamanla değişen ve etkileşimli ortamlarda karar alma süreçleri için umut verici araçlar sunsa da, pratikte bazı önemli sınırlılıklarla karşı karşıya kalmaktadır. Bu sınırlılıklar, özellikle yüksek boyutlu durum ve aksiyon uzaylarında gözlemlenmekte, çok sayıda parametreye ve sürekli kontrol gereksinimlerine sahip sistemlerde RL'in performansını sınırlandırmaktadır (Lilicrap ve ark., 2015). Bu durum, iletişim sistemlerinin giderek daha karmaşık hâle geldiği ve çoklu etkileşimlerin yoğunlaştığı V2X ortamlarında alternatif yöntem arayışlarını beraberinde getirmiştir. Bu arayışın en önemli sonuçlarından biri, DRL olarak literatüre giren yöntemlerdir. DRL, klasik RL'in temel yapısını korumakla birlikte, değer fonksiyonları ve politika öğrenimini temsil etmek için derin sinir ağlarını (deep neural networks, DNN) kullanır (Sze ve ark., 2017). Bu sayede çok geniş durum ve aksiyon uzaylarında genelleme yapabilme, gürültülü gözlemleri işleyebilme ve sürekli değerler içeren kontrol problemlerine çözüm üretebilme kapasitesine sahiptir.

V2X ortamları için kritik olan kontrol görevlerinin çoğu, doğası gereği sürekli aksiyon uzayına sahiptir. Örneğin; bir aracın ne kadar iletim gücü kullanması gerektiği, hangi kanal üzerinden ne yoğunlukta veri göndereceği veya iletişimi hangi modla gerçekleştireceği gibi kararlar, sabit ve ayrık değerlerden çok, sürekli aralıklar üzerinde alınan kararları ifade eder. Klasik RL algoritmaları (örneğin Q-learning, SARSA vb.) ayrık aksiyon alanlarına uygun yapıdadır ve bu nedenle sürekli kontrol gerektiren uygulamalarda doğrudan uygulanabilir değildir (Jang ve ark., 2019). Bu eksiklik, DRL algoritmalarını V2X sistemleri için tercih edilen bir yaklaşıma dönüştürmüştür.

DRL'nin sunduğu bu avantajlar doğrultusunda geliştirilen bazı algoritmalar, iletişim alanında öne çıkmıştır. Özellikle derin Q-ağları (deep Q-networks, DQN)

algoritması, büyük durum uzaylarında ayrık aksiyonlara sahip problemleri çözmek için oldukça etkilidir (Huang, 2020). DQN'nin sınırlı kaldığı nokta, sürekli aksiyon uzayıdır. V2X gibi ağlarda iletim gücü, zamanlama, frekans tahsisi gibi birçok parametre sürekli değişkenlerden oluştuğunda, DQN yapısı bu değerleri doğrudan işleyemez. Bu tür senaryolarda politika gradyanlı (policy gradient) yaklaşımlar ve aktör kritik mimarileri devreye girmektedir.

Bu bağlamda, derin deterministik politika gradyanı (deep deterministic policy gradient, DDPG) algoritması, DRL dünyasında sürekli aksiyon uzayları için en çok tercih edilen yöntemlerden biri olarak kabul edilmektedir. DDPG, politika temelli öğrenmeyi değer fonksiyonları ile birlikte optimize eden aktör-kritik yapısını kullanır. Burada "aktör" ajanı aksiyon üretirken, "kritik" bu aksiyonun ne kadar faydalı olduğunu değerlendirir. Böylece öğrenme süreci daha dengeli ve hızlı gerçekleşir. V2X sistemlerinde DDPG, özellikle iletim gücü ayarlama, kanal tahsisi, mod seçimi gibi sürekli kontrol problemleri için oldukça uygundur. DDPG'nin avantajı yalnızca sürekli uzayları işleyebilmesinde değil, aynı zamanda politika dışı yapısıyla geçmiş deneyimleri tekrar kullanarak öğrenme sürecini hızlandırmasında da yatmaktadır. Bu özellik, simülasyon ortamlarında veri üretiminin sınırlı olduğu veya fiziksel sistemlerde deneme-yanılma maliyetinin yüksek olduğu durumlarda ciddi avantaj sağlar. V2X sistemlerinde bu tür koşullar oldukça yaygın olduğundan, DDPG ve benzeri algoritmalar bu alanda giderek daha fazla ilgi görmektedir. DDPG, karar süreçlerinin yüksek zaman çözünürlüğü ile ele alınmasını da mümkün kılar. Yani bir araç, milisaniyeler içinde onlarca karar alabilir ve bu kararlar dinamik çevre koşullarına uygun olarak sürekli güncellenebilir. Bu da, AoI, enerji verimliliği ve güvenli haberleşme gibi performans kriterlerinde belirgin iyileştirmeler sağlar. Nitekim yapılan çalışmalar, DDPG tabanlı kontrol sistemlerinin klasik yöntemlere göre daha düşük gecikme, daha güncel veri iletimi ve daha verimli kaynak kullanımını sağladığını ortaya koymuştur (Parvini ve ark., 2023).

Ancak tüm bu avantajlarına rağmen, DDPG yalnızca tek ajanlı (single-agent) yapılar için tasarlanmıştır. Yani yalnızca bir aracın kendi kararlarını optimize etmesi hedeflenir. Oysa V2X sistemleri doğası gereği çoklu araçlar, etkileşimli bağlantılar ve kooperatif karar alma mekanizmaları içerir. Bu nedenle, sistemin daha gerçekçi biçimde modellenebilmesi için birden fazla ajanın aynı anda öğrenme ve karar alma sürecine dâhil olduğu çoklu ajan derin pekiştirmeli öğrenme (multi-agent deep reinforcement

learning, MADRL) yaklaşımlarına ihtiyaç duyulmaktadır (Zhang ve ark., 2020). Bu kapsamda bir sonraki bölümde ayrıntılı olarak ele alınacak olan MADDPG algoritması, çok ajanlı ortamlar için DDPG yapısını genişleterek V2X sistemlerine doğrudan uygulanabilir hâle getirmektedir.

### 2.2.5. Çok Ajanlı Ortamlar İçin MADRL Yaklaşımı

V2X sistemleri, doğası gereği çok sayıda bağımsız karar vericinin (araçların) aynı fiziksel ve mantıksal iletişim alanını paylaştığı, yüksek derecede etkileşimli ve dağıtık yapılardır. Her araç, çevresel farkındalık oluşturmak, yol koşullarını algılamak ve güvenli sürüş kararları almak amacıyla hem diğer araçlarla hem de altyapı elemanlarıyla eş zamanlı olarak iletişim kurmak zorundadır. Bu nedenle, V2X iletişim ağı sadece bireysel ajanların performansına değil, çoklu etmenlerin (multi-agent) birbirleriyle olan etkileşimine de doğrudan bağlıdır. Bu durum, klasik tek ajanlı RL/DRL yaklaşımlarını yetersiz kılmakta ve MADRL yöntemlerini gerekli hâle getirmektedir.

Tek ajanlı sistemlerde her ajan yalnızca kendi gözlemlerine dayanarak karar verir ve çevredeki diğer ajanların davranışlarını doğrudan dikkate almaz. Ancak V2X gibi kooperatif ya da rekabetçi etkileşimlerin yoğun olduğu ortamlarda, bir aracın aldığı karar sadece kendi performansını değil, aynı zamanda diğer araçların da haberleşme kalitesini, kaynaklara erişimini ve bilgi güncelliğini etkileyebilir. Örneğin bir araç iletim gücünü artırarak daha uzak araçlara sinyal göndermeyi hedeflediğinde, bu karar aynı kanalı kullanan başka araçların sinyal kalitesini bozabilir. Bu nedenle, ajanlar arasında ortak bağlamı paylaşan, ancak dağıtık karar alabilen bir yapının oluşturulması gerekmektedir.

Bu bağlamda MADRL yaklaşımları, birden fazla ajanın aynı çevreyi gözlemlediği ve kendi yerel bilgileri doğrultusunda hareket ettiği öğrenme ortamlarını temsil eder. Her bir ajan, kendi politika fonksiyonunu öğrenir; ancak diğer ajanların politikaları da çevrenin evrimi üzerinde belirleyici olduğundan, sistemin öğrenme süreci oldukça karmaşık bir hâl alır. Bu yapı içerisinde her bir ajanın aksiyonları yalnızca kendi bireysel ödülünü değil, aynı zamanda diğer ajanların ödülleri de etkileyebileceğinden, öğrenme sürecinde çoklu etki, karşılıklı bağımlılık ve koordinasyon ihtiyacı ortaya çıkar.

Bu karmaşık öğrenme ortamları için geliştirilen en başarılı algoritmalarından biri olan MADDPG, DDPG'nin temel yapısını çok etmenli yapılara uyarlayarak hem bireysel öğrenmeyi hem de ortak sistem optimizasyonunu mümkün kılar (Parvini ve ark., 2021). MADDPG'de her ajan kendi yerel gözlemine dayanarak aksiyon üretirken, bu aksiyonların genel sistem üzerindeki etkisi merkezi bir "kritik" tarafından değerlendirilir. Yani öğrenme süreci dağıtık aksiyon alma ile birlikte merkezi değerlendirme (centralized training with decentralized execution) prensibine dayanır. Bu yaklaşım sayesinde, her ajan gerçek zamanlı olarak bağımsız çalışabilirken, sistem eğitimi sırasında diğer ajanların davranışları da dikkate alınarak daha koordineli politika öğrenimi sağlanır.

MADDPG'nin V2X sistemlerine uygulanabilirliği oldukça yüksektir. Her araç bir ajan olarak ele alınabilir ve araçlar hem birbirleriyle (örneğin platoon içi haberleşme), hem altyapı (örneğin RSU), hem de HAPS gibi hücre dışı platformlar ile iletişim kurarak ortak görevleri gerçekleştirebilir. Araçların iletim gücünü ayarlaması, kanal seçiminde bulunması, hangi mod üzerinden iletişim kuracağına karar vermesi gibi işlemler, ajanların aksiyon alanlarını oluşturur. Bu aksiyonların optimize edilmesi, sadece bireysel bağlantı kalitesini değil, aynı zamanda sistemdeki genel AoI ve enerji verimliliğini de artırır.

MADDPG'nin önemli bir avantajı da sistemdeki heterojen hedefleri aynı anda optimize edebilmesidir. Örneğin bazı ajanlar AoI'yı düşürmeyi, bazıları enerji tüketimini azaltmayı, bazıları ise kanal yoğunluğunu azaltmayı hedefleyebilir. Merkezi kritik yapısı, tüm bu metrikleri hesaba katarak, ajanların öğrenmesini yönlendirebilir. Böylece çok boyutlu optimizasyon problemleri, tekil hedef fonksiyonlarına indirgenmeden çözülmüş olur. MADDPG, V2X ortamlarında karşılaşılan kaynak yönetimi, mod seçimi, spektrum paylaşımı ve bilgi güncelliği gibi çok yönlü problemleri aynı çatı altında ele alabilen güçlü bir öğrenme altyapısı sunmaktadır. Ajanların hem yerel hem de kolektif bilgiyi kullanarak öğrenmesi, iletişim sistemlerinin daha çevik, güvenli ve adaptif hâle gelmesini sağlar. Bu bağlamda, bu tez çalışmasında da MADDPG algoritması tercih edilerek, HAPS destekli V2X sistemlerinde bilgi güncelliği merkezli kaynak yönetimi hedeflenmiştir.

### 3. MATERYAL VE YÖNTEM

Bu bölümde, tez kapsamında ele alınan problem alanı ile doğrudan ilişkili olan temel teknolojiler, kavramsal çerçeveler ve çözüm yaklaşımları sistematik bir şekilde incelenmektedir. Hem çalışmanın dayandığı teknik temeller ortaya koyulmakta hem de mevcut yaklaşımların sınırları analiz edilmektedir.

İlk olarak V2X iletişim sistemlerinin mimarileri ve teknik altyapıları ele alınacak, ardından AoI kavramının sistem performansı açısından neden kritik olduğu açıklanacaktır. Daha sonra HAPS sistemlerinin V2X ortamlarına entegrasyonu incelenecek, kaynak yönetimi stratejileri değerlendirilerek mevcut yöntemlerin avantajları ve eksiklikleri tartışılacaktır. Son olarak, AI tabanlı karar sistemleri, özellikle DRL ve MADRL, HAPS destekli V2X sistemlerine yönelik çözüm potansiyelleri bağlamında ele alınacaktır.

#### 3.1. V2X İletişim Teknolojileri ve Mimarileri

V2X sistemleri, geleceğin akıllı ulaşım altyapılarının temelini oluşturmaktadır. Bu sistemler, yalnızca bireysel araçlar arasında değil; aynı zamanda altyapı, ağ servis sağlayıcıları, yayalar ve hava tabanlı haberleşme birimleri gibi çok çeşitli ağ bileşenleri arasında gerçek zamanlı bilgi alışverişini mümkün kılar. Bu sayede, trafik güvenliğinden enerji verimliliğine, sürüş konforundan otonom karar sistemlerine kadar pek çok alanda dönüşüm yaratılmaktadır. V2X mimarisi, araçların çevreleriyle yüksek hızda ve düşük gecikmeyle veri alışverişinde bulunmasını sağlayarak, çevresel farkındalıklarını artırmakta ve otonom sistemlerin güvenliğini doğrudan etkilemektedir. Bu bağlamda, V2X sistemlerinin teknik temelleri yalnızca bir iletişim protokolü olmaktan çıkmış; çok katmanlı, heterojen ve bağlama duyarlı bir sistem mimarisine dönüşmüştür.

V2X haberleşmesi uzun yıllar boyunca sabit altyapıya bağlı olarak şekillenmiş, çoğunlukla özel kısa menzilli iletişimler (dedicated short-range communication, DSRC) protokolüne dayanan çözümlerle yürütülmüştür. DSRC, düşük gecikme süresi ve doğrudan bağlantı imkânı sunması sebebiyle özellikle erken dönem güvenlik uygulamalarında tercih edilmiştir. Ancak bu yapının kapsama alanı dardır, kanal yoğunluğuna karşı duyarlıdır ve ölçeklenebilirliği sınırlıdır (Kenney ve ark., 2011). Bu

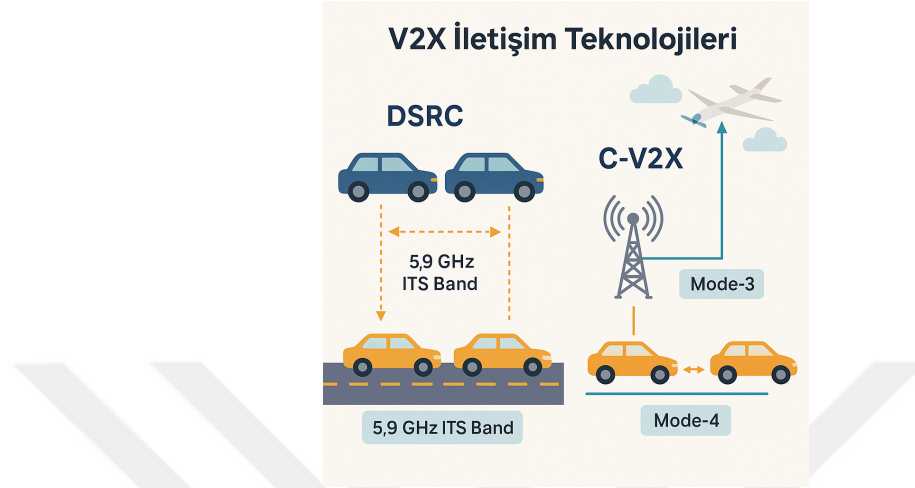
nedenlerle, DSRC'ye alternatif olarak hücreli V2X (cellular V2X, C-V2X) yapısı geliştirilmiş ve daha geniş kapsama alanı, altyapı destekli bağlantı ve yüksek bant genişliği gibi avantajlarla sistemin sınırları genişletmiştir (Liu ve ark., 2020). C-V2X, hem doğrudan araçlar arası iletişimi hem de baz istasyonu destekli geniş alan bağlantılarını aynı mimari altında topladığı için, hem yerel hem küresel veri iletişimini optimize edebilen bir yapı sunmaktadır. Ancak bu sistem dahi, hızlı değişen ağ topolojisi, ani kanal bozulmaları, çok sayıda kullanıcının eş zamanlı veri iletme ihtiyacı gibi sorunlar karşısında mevcut yöntemlerle sürdürülebilir çözümler sunmakta zorlanmaktadır.

Bu kapsamda 5G-V2X ve sonrasında şekillenmeye başlayan 6G-V2X çalışmaları, yalnızca iletişim hızını değil; bağlamsal farkındalığı, karar desteğini ve sistem zekâsını da incelemeye başlamıştır. 6G ile birlikte ağ mimarilerinde görülen evrim, veri işleme ve depolama kaynaklarının da entegre edildiği bir yapı oluşturarak, ağın karar alma kapasitesini artırmayı hedeflemektedir (Alwarafy ve ark., 2022). Bu paradigma değişimi, klasik ağ katmanları olan fiziksel, bağlantı, ağ ve uygulama katmanlarının sınırlarını daha geçirgen hâle getirmekte ve iletişim-bilişim-depolama yakınsaması gibi yeni kavramları literatüre taşımaktadır. Artık bir ağ düğümü yalnızca veri iletmekle kalmayıp; aynı zamanda çevreden gelen veriyi anlamlandırma, ön işleme tabii tutma ve gerektiğinde karar verici rolü üstlenme potansiyeline sahiptir. Bu da V2X sistemlerini, geleneksel ağlardan farklı olarak dinamik, dağıtık ve bağlama duyarlı yapılara dönüştürmektedir.

Bu bağlamda, V2X sistemlerinin günümüzde karşı karşıya kaldığı en büyük zorluklardan biri, çok katmanlı ve çok etmenli ortamda kaynakların nasıl yönetileceği sorusudur. Her araç hem alıcı hem gönderici olarak görev yapmakta; aynı anda birden fazla bağlantıya ihtiyaç duymakta ve çevredeki diğer araçların iletişim tercihleriyle doğrudan etkileşim hâlinde olmaktadır (Parvini ve ark., 2023). Bu durum, sistemin yalnızca bireysel performansla değil, kolektif denge ve etkileşimli kararlar ile sürdürülebilir hâle getirilebileceğini göstermektedir. Ancak bu dengeyi sağlamak, geleneksel sabit kurallara dayalı karar sistemleriyle mümkün değildir. Çevreye duyarsız, önceden belirlenmiş eşiklere bağlı çalışan yapılar, yüksek hareketlilik ve kanal değişkenliği altında hem bilgi güncelliğini bozar hem de sistem verimliliğini düşürür.

Dolayısıyla V2X sistemlerinin gelişimi, sadece altyapının genişlemesiyle değil; aynı zamanda bu altyapının nasıl yönetileceği, öğrenip uyum sağlayabileceği, ve gerçek

zamanlı kararlar alabileceği ile doğrudan ilişkilidir. Bu yönüyle, mimari tasarımlar artık donanım ve bağlantı katmanlarıyla sınırlı kalmamakta; öğrenen sistemler, dağıtık karar mekanizmaları ve AI destekli haberleşme yönetimi gibi yazılımsal bileşenlerin entegrasyonunu da gerektirmektedir (Zhang ve ark., 2024).



Şekil 3.1. DSRC ve C-V2X yapılarına ilişkin karşılaştırmalı mimari.

V2X sistemlerinin teknik olarak gerçekleştirilmesi için Şekil 3.1'de görülen iki temel protokol hattı öne çıkmaktadır: DSRC ve C-V2X. DSRC, 5,9 GHz ITS bandında çalışarak doğrudan araçlar arası haberleşmeye olanak tanır (Arena ve ark., 2020). Fakat fiziksel katmanda kullandığı ortogonal frekans bölmeli çoklama (Orthogonal frequency-division multiplexing, OFDM) tekniği yoğun trafik senaryolarında çarpışmalara ve gecikmelere yol açabilmektedir. Bu sebeple, daha ileri düzey hizmet kalitesi (quality of service, QoS) hedeflerini gerçekleştirmek için üçüncü nesil ortaklık projesi (third generation partnership project, 3GPP) tarafından tanımlanan C-V2X çözümleri, özellikle uzun vadeli evrim versiyon 14 (long term evolution release 14, LTE-Rel.14) ve sonrasında beşinci nesil (fifth generation, 5G) yeni radyo V2X (new-radio V2X, NR-V2X) ile gelişim göstermiştir (Gomez-Barquero ve ark., 2020).

C-V2X teknolojisi, Mode-3 ve Mode-4 olmak üzere iki temel haberleşme kipinde çalışır. Mode-3, baz istasyonu tarafından kaynakların merkezi olarak tahsis edildiği yapıyı temsil ederken; Mode-4, baz istasyonu bağlantısının olmadığı durumlarda, araçların kaynakları özerk biçimde seçmesini sağlar (Saad ve ark., 2025). Bu bağlamda, V2X iletişim sistemlerinin mimarisi ve kullanılan haberleşme protokolleri, sistemin genel performansını doğrudan etkilemektedir. DSRC, araçlar arasında doğrudan kısa menzilli iletişim sağlar ve düşük gecikme avantajı sunar; ancak

özellikle yoğun trafik senaryolarında kanal çakışmalarına karşı hassastır. Buna karşın, C-V2X mimarisi daha geniş kapsama alanı ve esnek kaynak yönetimi sağlamaktadır. Mode-3 yapılandırmasında baz istasyonu kaynakları planlarken, Mode-4 senaryosunda araçlar kendi aralarında kaynak paylaşımı yapar. Ayrıca 6G ve ötesine yönelik vizyonlar doğrultusunda NTN gibi yapılar, haberleşme sürekliliği ve düşük gecikmeli bağlantı desteği açısından V2X sistemlerine entegre edilebilir. Bu mimariler, özellikle altyapı yetersizliği olan kırsal bölgelerde veya afet durumlarında kritik iletişim sürekliliği sağlama potansiyeline sahiptir.

### 3.2. AoI Kavramı

AoI, temel olarak bir bilginin üretildiği andan itibaren geçen sürenin, bu bilgi hâlâ kullanılmaya devam ediyorsa alıcı açısından ne kadar “yaşlı” olduğunu ölçer. Matematiksel olarak bu, her veri paketi için bir zaman damgası tutularak hesaplanır. Örneğin, bir araç fren yaptığında bu durumu paylaşan veri paketi hemen alıcıya ulaşsa dahi, eğer bu paket bir önceki güncellenmeden sonra geç iletilmişse veya araya zaman farkı girmişse, alıcı tarafındaki AoI değeri yüksek kalır. Buradaki kritik fark, gecikme metriğinin bir iletim süresini ölçmesi; AoI’nın ise sistemin o andaki bilgiye ne kadar süredir güncelleme almadığını hesaplamasıdır. Gecikme sifıra yakın olabilir; ancak eğer paketler arasındaki güncellemeler seyrekse ya da düzenli değilse AoI yüksektir. Bu fark, özellikle V2X gibi zamanın kritik olduğu sistemlerde, bilgi doğruluğu ile birlikte bilgi tazeliğini sistem performansının merkezine yerleştirmektedir.

Bu nedenle AoI, yalnızca ağın taşıma kapasitesi ile ilgili değil; aynı zamanda bilginin operasyonel anlamda işe yararlılığı ile doğrudan ilgilidir. Örneğin bir aracın hız bilgisi saniyede bir güncellenirken, o aracın frenleme ya da yön değiştirme eylemleri milisaniyelik aralıklarla gerçekleşebilir. Bu durumda saniyelik bir bilgi akışı, sistemin otonom karar mekanizması için yetersiz kalacaktır. Araç, aslında mevcut hız bilgisini doğru alıyor olsa bile, AoI yüksekse, bu bilgi artık sistem tarafından “eski” kabul edilmeli ve güvenlik kararlarında kullanılmamalıdır. İşte bu noktada AoI, veri iletim performansından ayrılarak doğrudan karar kalitesini etkileyen bir sistem metriği hâline gelmektedir.

Literatürde, AoI’nın bu yönüyle yalnızca iletim süreçlerini değil; ağ tasarımı, kaynak yönetimi ve önceliklendirme stratejilerini de etkilediği gösterilmiştir. Özellikle

V2X sistemlerinde sıkça karşılaşılan kanal yoğunluğu, frekans çakışmaları ve zamanlama çelişkileri gibi problemler, paketlerin sistemdeki yaşlarını doğrudan etkileyerek AoI'yı artırmakta, bu da bilgiye dayalı kararların doğruluğunu azaltmaktadır (Maglogiannis ve ark., 2021). Dolayısıyla bilgi güncelliğinin korunması yalnızca ağ üzerindeki veri akışını değil, aynı zamanda kaynak tahsisinin nasıl yapılacağına dair kararları da doğrudan yönlendirmektedir.

Buna ek olarak, V2X sistemlerinde yalnızca birden fazla araç değil, aynı zamanda RSU'lar, merkezi sistemler ve son dönemde HAPS gibi farklı haberleşme yolları da gündemde olduğu için, her bir veri için sadece ne zaman gönderileceği değil, hangi yoldan gönderileceği, hangi öncelikte iletileceği ve hangi güç düzeyinde aktarılacağı gibi çok sayıda değişkenin birlikte optimize edilmesi gerekmektedir (Alwarafy ve ark., 2022). Bu karmaşık karar yapısında, AoI yalnızca bir izleme metriği değil, aynı zamanda sistemsel optimizasyon hedefi olarak da kullanılabilir hâle gelmiştir.

V2X sistemlerinde AoI'nın düşük tutulması, bilgiye dayalı karar süreçlerinin güvenilirliğini artırmakta, sistem tepkilerini hızlandırmakta ve özellikle otonom sürüş senaryolarında hayati önem taşımaktadır. Ancak AoI'nın optimize edilmesi, yalnızca gönderim sıklığını artırmak veya iletişim yollarını hızlandırmakla çözülebilecek bir problem değildir. Aksine, bu süreç çok boyutlu bir optimizasyon problemine dönüşmektedir. Örneğin veri paketlerinin sıklıkla gönderilmesi kısa vadede AoI'yı düşürebilir; fakat kanal tıkanıklığını ve çakışmaları artırarak gecikmelere ve paket kayıplarına neden olabilir. Benzer şekilde, iletim gücünün artırılması iletişimin başarımını artırabilirken, enerji tüketimini ve diğer araçlarla olan etkileşim yoğunluğunu da artırarak sistem dengesini bozabilir (Parvini ve ark., 2023).

Bu tür çelişkili hedefler nedeniyle AoI'nın optimize edilmesi, çok sayıda değişkenin dinamik olarak yönetildiği, çevreye duyarlı karar mekanizmaları gerektirir. Özellikle kanal kalitesi, trafik yoğunluğu, hareketlilik düzeyi, bağlantı tipi (V2V, V2I, V2H) ve veri önceliği gibi faktörlerin her an değiştiği bir ortamda, sabit eşik değerlerine dayalı kurallar yerine öğrenmeye dayalı, bağlama duyarlı sistemlerin kullanılması zorunludur. Bu bağlamda, son yıllarda DRL algoritmaları, V2X sistemlerinde AoI odaklı kontrol ve optimizasyon süreçlerinde etkili bir çözüm alanı sunmaya başlamıştır (Zhang ve Ark., 2020).

Çalışmamızda önerilen HAPS destekli MADDPG mimarisi, bu çok etmenli öğrenme ortamına yüksek irtifa platformlarının sunduğu düşük gecikmeli, sürekli bağlantılı, geniş kapsamlı iletişim desteğini entegre ederek AoI açısından belirgin performans kazanımları sunmaktadır.

### 3.3. HAPS Sistemleri ve V2X Entegrasyonu

Haberleşme altyapısının tüm dünyaya eşit şekilde yayılmadığı günümüz koşullarında, özellikle yüksek hareketlilik ve geniş alan kapsama gerektiren uygulamalar için NTN giderek daha fazla önem kazanmaktadır. Bu kapsamda HAPS, sabit kanatlı hava araçları veya balon benzeri platformlar aracılığıyla 20 ila 50 kilometre irtifada konuşlanarak, geniş kapsama alanı ve sürekli bağlantı sunabilen stratejik sistemler olarak öne çıkmaktadır. HAPS sistemleri, hem uyduların yüksek gecikme ve yüksek maliyet sorunlarına alternatif sunmakta hem de karasal ağların kapsama eksikliği yaşadığı bölgelerde köprü görevi üstlenmektedir.

Teknolojik olarak değerlendirildiğinde HAPS sistemleri genellikle güneş enerjisiyle çalışan, uzun süre havada kalabilen ve düşük maliyetli bakım gerektiren hafif hava taşıtlarıdır. Gelişmiş yönlendirme sistemleri sayesinde belirli bir alan üzerinde sabit kalabilir ya da dinamik olarak yeniden konumlandırılabilirler. Bu özellik, farklı trafik yoğunluğu senaryolarında HAPS'in ağ talebine göre konumlandırılmasına ve ağ yükünün dengelenmesine olanak sağlar. Ayrıca, yeniden konuşlandırılabilir yapıları sayesinde afet gibi olağanüstü durumlarda çöken karasal altyapıların yerine kısa sürede iletişim hizmeti sunabilirler.

HAPS'in bir diğer avantajı, uydulara kıyasla çok daha düşük gecikme süreleri ile çalışmasıdır. Örneğin jeostatik yörünge (geostationary orbit, GEO) uyduları 35.000 km uzaklıkta ve yaklaşık 600 ms'lik gecikme ile çalışırken, HAPS sistemleri birkaç 10 ms civarında bir uçtan uca gecikme sunabilir. Bu özellik, özellikle AoI kritik uygulamalarda, HAPS'ı tercih edilir bir çözüm hâline getirmektedir. Aynı zamanda kapsama alanı açısından değerlendirildiğinde, tek bir HAPS cihazı ortalama 50 ila 200 kilometrelik yarıçapta bir bölgeye hizmet verebilmektedir. Bu değer, özellikle karasal baz istasyonlarının yaygın olmadığı kırsallarda ve otoyol bölgelerinde tek başına bile etkili bir iletişim altyapısı oluşturmak için yeterlidir.

Son yıllarda 3GPP standartlarının NTN çalışmalarında HAPS sistemlerine özel olarak yer vermesi, bu platformların yalnızca araştırma laboratuvarlarının değil; aynı zamanda uluslararası telekomünikasyon otoritelerinin de radarında olduğunu göstermektedir. Örneğin 3GPP Release 17 ile başlayan NTN standardizasyon süreci, HAPS'ı hem bağımsız iletişim istasyonu hem de hücreli ağların genişletici unsuru olarak tanımlamaktadır (Saad ve ark., 2024). Bu, HAPS'ın 6G öncesi geçiş dönemi ve 6G sonrası hiper bağlantılı mimariler için aktif rol oynayacağını kanıtlar niteliktedir.

Bu teknik üstünlükleri göz önüne alındığında, HAPS sistemlerinin sadece kırsal alanlarda acil iletişim çözümleri sunan alternatifler değil; aynı zamanda gelişmiş haberleşme mimarilerinin kalıcı bileşenleri olmaya aday olduğu açıkça görülmektedir. Özellikle sabit altyapı kurulumunun zor, maliyetli ya da yetersiz olduğu senaryolarda, HAPS sistemleri yüksek erişilebilirlik, düşük bakım gereksinimi, geniş kapsama ve ağ esnekliği ile öne çıkmaktadır. Bu özellikler, V2X gibi anlık veri iletiminin, yüksek mobilitenin ve ağ adaptifliğinin bir arada bulunduğu uygulamalarda HAPS'ı çok değerli bir aday hâline getirmektedir.

Günümüzde HAPS sistemlerinin gelişimi yalnızca teorik kavramlarla sınırlı kalmamış, çeşitli ticari ve araştırma projeleriyle uygulama alanı bulmuştur. Örneğin, Google tarafından geliştirilen Project Loon, stratosferik balonlar aracılığıyla geniş alanlara internet erişimi sağlamayı hedeflemiş ve HAPS konseptinin uygulanabilirliğini küresel ölçekte göstermiştir. Öte yandan, Airbus Zephyr gibi sabit kanatlı, güneş enerjisiyle çalışan hava platformları da hem ticari hem de askeri haberleşme çözümleri kapsamında test edilmektedir. Bu örnekler, HAPS teknolojilerinin yalnızca kırsal erişim alternatifi değil, aynı zamanda çok yönlü haberleşme altyapısı bileşeni hâline geldiğini ortaya koymaktadır.

HAPS, farklı yapısal türleri ve görev profilleri ile V2X iletişim altyapısına entegre edilebilecek geniş bir yelpazede çözümler sunmaktadır. Şekil 3.2'te yatay düzlemde görselleştirilen bu platformlar; balon tabanlı sistemler, sabit kanatlı güneş enerjili hava araçları ve hibrit ileri teknoloji platformları olarak sınıflandırılabilir. Bu sınıflandırma, hem teknik kapasitelerine hem de kullanım senaryolarına göre belirlenmiştir. Bu uygulamalar sadece teoride kalmamış, uygulanmaya başlanmıştır. Google tarafında geliştirilen Project Loon ve Airbus tarafından geliştirilen Airbus Zephyr bunların gerçek dünya örneklerindedir.



Şekil 3.2. HAPS örnekleri

Standart V2X sistemlerinde iletişim modları genellikle karasal altyapıya göre seçilmekte ve kanal koşulları bu tercihlere bağlı kalmaktadır. Oysa HAPS, kanal ortamı yoğunlaştığında veya belirli bağlantılar kopma riski taşıdığı anda alternatif bağlantı yolları sağlayarak sistemin esnekliğini artırır. Bu esneklik, özellikle çok etmenli karar ortamlarında, ajanların (araçların) bağlantı stratejilerini bağlamsal olarak optimize etmesine olanak tanır. Her araç kendi konumuna, kanal kalitesine ve enerji seviyesine göre karar alırken, HAPS destekli alternatif bağlantı yollarını kullanarak sistemin genel performansına zarar vermeden bireysel optimizasyon gerçekleştirebilir. Bu durum, hem kaynak paylaşımındaki çatışmaları azaltmakta, hem de ağın öğrenme sürecinde dengeli politika gelişimine katkı sunmaktadır (Kurt ve ark., 2021). HAPS destekli haberleşme, yalnızca iletişim sürekliliği açısından değil; aynı zamanda ölçeklenebilirlik ve yük dengeleme açısından da sistem mimarisine katkı sağlar. Özellikle şehir içi yüksek yoğunluklu trafik bölgelerinde, karasal baz istasyonları hızla aşırı yük altına girebilirken, HAPS bu yükün bir kısmını üzerine alarak ağın tıkanmasını önleyebilir. Aynı zamanda trafik yoğunluğuna bağlı olarak HAPS platformlarının hareket ettirilmesi ya da görev bölgesinin değiştirilmesi sayesinde, sistem dinamik trafik koşullarına adaptif biçimde yanıt verebilir. Bu tür bir dinamik yeniden konuşlanma yeteneği, yalnızca kapsama alanı yönetiminde değil; aynı zamanda ağ öğrenme politikalarının güncellenmesinde ve sistem davranışlarının iyileştirilmesinde de stratejik fayda sağlar.

Tüm bu avantajlara rağmen, HAPS destekli V2X mimarileri hâlâ akademik literatürde sınırlı sayıda çalışmada yer bulmakta, var olan araştırmalar ise genellikle

sistem düzeyinde tasarım önerileri veya kanal modeli analizleri ile sınırlı kalmaktadır. Oysa bilgi güncelliği gibi gerçek zamanlı metriklerin, çok etmenli öğrenme algoritmalarıyla birlikte HAPS destekli V2X ağlarında optimize edilmesi, literatürde henüz yeterince işlenmemiş açık bir araştırma boşluğu sunmaktadır. Bu bağlamda, HAPS destekli sistemlere V2X entegrasyonunun yalnızca teknik bir bağlantı birleşimi değil; aynı zamanda bir sistem zekâsı dönüşümü olduğu göz önünde bulundurulmalıdır.

### 3.4. V2X'te Kaynak Yönetimi Yöntemleri

V2X sistemleri, çok sayıda hareketli birimin (araçların) sınırlı sayıda ağ kaynağını eş zamanlı olarak kullanmak zorunda olduğu, yüksek yoğunluklu ve dağıtık yapıli haberleşme ortamlarıdır. Bu tür ortamlarda iletişim sürekliliğini ve hizmet kalitesini sürdürülebilir biçimde sağlayabilmek için ağ kaynaklarının (zaman aralıkları, frekans blokları, kanal bant genişliği, iletim gücü ve bağlantı yolları gibi) verimli bir biçimde yönetilmesi gerekir. Bu gereksinim, V2X mimarilerinde kaynak yönetimini yalnızca teknik bir görev olmaktan çıkarıp, stratejik bir optimizasyon problemine dönüştürmektedir.

V2X sistemlerinin sahip olduğu temel karakteristikler (yüksek mobilite, düzensiz bağlantı topolojisi, heterojen iletişim modları, görev önceliklerinin değişkenliği ve gerçek zamanlı karar alma zorunluluğu) kaynakların geleneksel yöntemlerle paylaşılmasını son derece güç hâle getirir. Bu sistemlerde kaynaklar yalnızca eş zamanlı olarak değil, aynı zamanda koordineli ve bağlama duyarlı biçimde tahsis edilmelidir. Aksi takdirde, bazı araçların veri iletiminde aşırı avantaj kazanması, bazılarının ise bilgi akışından dışlanması gibi dengesizlikler oluşur. Bu da ağ üzerinde hem AoI bozulmasına hem de kaynak israfına neden olur.

Bu bölümde, klasik yöntemlerin neden günümüz V2X yapıları için yetersiz kaldığı, modern yöntemlerin ne tür avantajlar sunduğu ve ileri düzey öğrenme temelli yapıların neden kaçınılmaz olduğu açıkça ortaya konulacaktır. Böylece, tez kapsamında önerilen HAPS destekli ve AoI odaklı kaynak yönetimi çözümünün gerekçesi hem teorik hem de pratik düzeyde temellendirilmiş olacaktır.

### 3.4.1. Geleneksel Kaynak Yönetimi Yöntemleri ve Sınırları

V2X sistemlerinin erken dönemlerinde, ağ kaynaklarının yönetimi görece daha az karmaşık topolojilere, sabit trafik yüklerine ve sınırlı mobiliteye sahip ortamlarda uygulanmak üzere tasarlanmıştır (Seid ve ark., 2023). Bu dönemde, ağ üzerindeki haberleşme birimleri az sayıda ve çoğunlukla sabit olduğundan, önceden tanımlanmış sabit kurallar, zaman dilimi tabanlı erişim, round-robin algoritmaları ve öncelik temelli sabit politika setleri kaynak tahsisi için yeterli görünmekteydi (Gyawali ve ark., 2020). Ancak V2X sistemleri zamanla daha hareketli, daha yoğun, daha etkileşimli ve çok katmanlı hâle geldikçe, bu yöntemlerin sınırlılıkları belirginleşmiştir.

En temel geleneksel yöntemlerden biri olan sabit zamanlama (fixed scheduling), her araca belirli zaman dilimlerinde kanal erişimi tanır. Bu yapı, karışıklığı ve çakışmayı önlemek adına oldukça sade ve öngörülebilirdir. Ancak bu öngörülebilirlik, dinamik ortamlar için bir dezavantaj hâline gelir. Çünkü sistemdeki trafik yükü sabit değildir; araçlar sürekli hareket hâlinindedir ve kanal koşulları anlık olarak değişmektedir. Dolayısıyla sabit zaman dilimlerine dayalı erişim, bazı araçların kanal gereksinimlerini karşılayamazken, bazı araçların tahsis edilen süreden tam kapasiteyle faydalanmamasıyla kaynak israfına yol açar (Trabelsi ve ark., 2019).

Bununla bağlantılı bir diğer strateji olan round-robin zamanlama, sırayla kanal erişimi verir ve ayırım yapmadan tüm cihazların eşit haklara sahip olmasını sağlar. Bu yöntem adaletli görünse de, V2X sistemlerinde bilgi türleri eşit öneme sahip değildir. Örneğin çarpışma uyarısı paketi ile hava durumu güncellemesi aynı öncelikte değerlendirilemez. Bu gibi durumlarda round-robin stratejisi, kritik verilerin gecikmesine neden olarak güvenlik zafiyeti oluşturabilir. Aynı zamanda bu yöntem, yüksek veri üreten araçlar ile düşük veri üretenler arasında trafik asimetrisi yarattığı için, kanal verimliliğini düşürür (Avcil ve ark., 2024).

Yaygın olarak kullanılan bir diğer geleneksel yaklaşım da öncelik temelli erişim kontrolüdür. Bu yöntemi kullanan Choi ve ark. (2022)'nin çalışmasında, her veri tipi için önceden tanımlanmış öncelik seviyeleri atanır. Örneğin acil durum verileri yüksek öncelikli kabul edilirken, konfor amaçlı veriler düşük öncelik alır. Bu yapı ilk bakışta mantıklı görünse de, sistemin adaptasyon yeteneği olmadığından sorunlar ortaya çıkar. Çünkü öncelik seviyesi statik olarak belirlendiği için, o anki bağlama göre anlamlı

olmayabilir. Örneğin bir fren uyarısı boş bir yolda anlamsız olurken, yoğun trafikte hayat kurtarıcıdır. Ancak sabit öncelik politikası bu farkı gözetmez.

Geleneksel sistemlerde yaygın olarak tercih edilen bir başka yöntem de merkezî kaynak yönetimidir. Bu yapıda, ağın ortasında bulunan bir kontrol birimi (örneğin merkezi sunucu ya da baz istasyonu) tüm kararları verir. Her araç; kanal erişimi, iletim gücü ya da zaman dilimi taleplerini merkeze iletir ve karar oradan geri döner (Ernst ve ark., 2023). Teorik olarak bu yapı, ağın genel durumu hakkında en fazla bilgiye sahip olan bir birimin karar vermesi nedeniyle küresel optimizasyon için uygundur. Ancak V2X gibi yüksek mobiliteye ve yoğun veri trafiğine sahip ortamlarda bu yaklaşım gerçek zamanlılık, karar gecikmesi, hesaplama yükü ve tek hata noktası gibi nedenlerle pratikte ciddi zorluklar içerir. Ayrıca merkezî yapı, özellikle kırsal bölgelerde veya altyapı dışı senaryolarda tamamen işlevsiz kalabilir.

Tüm bu yöntemler, ağın önceden tanımlanmış ve sabit varsayımlara dayandığı koşullarda çalışabilirlik sunar. Ancak V2X sistemleri doğası gereği belirsiz, değişken, bağlama duyarlı ve karar etkileşimli ortamlardır. Dolayısıyla kaynak yönetiminin yalnızca “kime ne kadar kaynak verileceği” değil; aynı zamanda “kaynağın kimin için ne zaman ve neden kritik olduğu” bilgisiyle yapılması gerekir. Bu tür yapısal farkındalık ise sabit politikalarla mümkün değildir. Örneğin, ağda bağlantı yoğunluğu artarken sabit zamanlamaya devam etmek, AoI değerlerinin hızla yükselmesine; yüksek öncelikli verilere aşırı ayrıcalık tanımak ise bazı veri türlerinin sistemden tamamen dışlanmasına neden olabilir. Tüm bunlar, ağ dengesizliğine ve performans erozyonuna yol açar.

Geleneksel kaynak yönetimi yöntemleri, V2X gibi çok etmenli, bağlama duyarlı ve hareketli yapılar için yeterli değildir. Sistem, yalnızca kaynakları paylaşmakla kalmayıp, bağlama göre bu paylaşımı anlamlandırmak zorundadır. Bu bağlamda geleneksel yöntemler önemli bir başlangıç noktası sunmakla birlikte, günümüz V2X sistemlerinin ihtiyaç duyduğu esneklik, öğrenme kabiliyeti ve çevresel adaptasyonu sağlayamaz. Bu nedenle kaynak yönetimi anlayışının daha esnek, dağıtık ve dinamik yapılara evrilmesi kaçınılmaz hâle gelmiştir.

### 3.4.2. Güncel Kaynak Yönetimi Yöntemleri

V2X sistemlerinin gerçek zamanlı, yüksek mobiliteli ve çok katmanlı yapısı; sabit ve merkezi kaynak yönetimi stratejilerinin yerini, giderek daha fazla dinamik ve dağıtık kararlara dayalı sistemlere bırakmasını zorunlu hâle getirmiştir. Bu tür ortamların doğası gereği, ağ üzerindeki her etmenin – araç, RSU ya da merkezi birim fark etmeksizin – kendi durumsal farkındalığına ve karar alma kapasitesine sahip olması beklenir. Çünkü anlık kanal koşulları, bağlantı yoğunluğu, iletim öncelikleri ve hatta çevresel değişkenler (hava durumu, trafik durumu, hızlanma/decelere) gibi faktörler, klasik sistemlerde olduğu gibi önceden tanımlanmış politikalara değil, bağlamsal değerlendirmelere göre karar verilmesini gerektirir.

Bu ihtiyacın bir sonucu olarak geliştirilen dinamik zamanlama algoritmaları, sabit zaman dilimlerine dayalı eşit erişim yerine, kanalın anlık durumuna göre erişim pencereleri oluşturarak kaynakları daha verimli kullanmayı hedefler. Örneğin kanal kalitesi iyi olan bir araç, düşük hata oranı ile hızlı veri aktarımı sağlayabileceğinden, sistem ona o anda daha fazla zaman dilimi ayırabilir. Buna karşılık kötü kanal koşullarına sahip bir araç, ya beklemeye alınır ya da daha kısa bir zaman dilimiyle sınırlı tutulur. Bu yapı hem kaynak israfını azaltır hem de iletim başarımını yükseltir (Yuan ve ark., 2021). Ancak bu yapıların çoğu merkezi değerlendirmeye bağlı olduğu sürece, karar gecikmesi ve ölçeklenebilirlik problemleri devam eder.

Bu sınırlamaları aşmak için geliştirilen bir başka metodoloji ise dağıtık zamanlama ve erişim kontrolü protokolleridir. Bu yöntemlerde her araç veya ağ düğümü, kendi çevresel bilgisini (kendi konumu, kanal ölçümü, komşu trafiği gibi) kullanarak bağımsız karar verir (Gyawali ve ark., 2020). Son dönemde özellikle öne çıkan bir yaklaşım, QoS kaynak yönetimi sistemleridir. Bu sistemlerde her veri tipi, önceden belirlenmiş QoS hedeflerine sahiptir: örneğin gecikme eşiği, paket kayıp toleransı, minimum bant genişliği gibi. Ağ, bu hedeflere ulaşılmasını önceliklendirerek kaynak tahsisini yönetir (Saxena ve ark., 2025). Ancak V2X gibi dinamik yapılarda, bu hedeflerin sabit kalması her zaman anlamlı değildir. Örneğin bir çarpışma uyarısı normal trafikte düşük gecikmeyle iletilmelidir; ancak yoğun trafikte bu hedefin mutlak şekilde sağlanması mümkün olmayabilir. Bu durumlarda esnek politika ayarı ve öncelik yeniden düzenlemesi gerekir – ki klasik QoS sistemleri bunu çoğu zaman gerçekleştiremez.

Dinamik kaynak yönetimi sistemlerinin etkin çalışabilmesi için, ağdaki her karar vericinin hem kendi durumunu hem de çevresinin durumunu algılayabilmesi gerekir. Bu yapıdaki algoritmalar genellikle lokal bilgiye dayalı tahmin ve karar mekanizmaları içerir. Örneğin bir araç, çevresindeki 2-3 aracın kanal kullanımı hakkında kestirim yaparak, kendi iletim penceresini buna göre ayarlayabilir. Bu tür lokal koordinasyon, çok etmenli çatışmaları azaltırken; gecikme ve yeniden iletim ihtiyacını da düşürür. Ancak bu sistemlerin başarımı doğrudan tahmin doğruluğuna, bilgi paylaşım sıklığına ve öğrenme hızı gibi faktörlere bağlıdır. Ayrıca lokal kararlar zaman zaman küresel optimumdan sapmalara neden olabilir. Bu nedenle, lokal karar yapılarının bağlama göre güncellenebilen esnek politikalara sahip olması gerekir. Literatürde bu tarz yaklaşımları destekleyen birçok model önerilmiştir. Örneğin oyun-teorik kaynak tahsisi (game-theoretic resource allocation) yapıları, her aracı bir oyuncu olarak modelleyerek, kaynak paylaşımını rekabet/iş birliği stratejileri üzerinden optimize eder (Nair ve ark., 2024).

Kısacası, dinamik ve dağıtık yöntemler geleneksel yöntemlere göre önemli avantajlar sunsa da, bunların büyük bir kısmı öğrenme yeteneğinden yoksun, tahmine dayalı, statik kural setlerine bağlı çalışır. Bu da sistemin uzun vadede değişen bağlama adapte olmasını engeller. Bu noktada çözüm, sadece dağıtık kararlar almak değil; bu kararların özerk ajanlar tarafından bağlam farkındalığıyla öğrenilerek verilmesini sağlamaktır. Dağıtık ve dinamik yaklaşımlar bu açığı bir nebze kapatsa da, günümüzde V2X sistemlerinin karmaşıklığı artık yalnızca "kimin ne zaman erişim sağlayacağı" sorusunu değil, aynı zamanda bu kararın hangi bağlamda verileceğini de temel bir sorun hâline getirmiştir.

Bu sorunu çözmek için öğrenebilen sistem mimarileri önerilmiştir. Özellikle RL temelli yapılar, bir ajan ile ortam arasındaki etkileşimden öğrenerek kendi politikasını geliştirir. Bu sayede sistem, zaman içinde çevre koşullarına adapte olmayı, hatalardan ders çıkarmayı ve uzun vadeli ödülleri maksimize etmeyi öğrenir. Klasik RL algoritmaları çoğunlukla tek ajanlı ortamlarda etkin çalışır; oysa V2X gibi sistemler birden çok bağımsız ve etkileşimli ajanın aynı anda karar aldığı ortamlardır. Bu nedenle klasik RL'in doğrudan uygulanması performans sorunlarına ve politika çakışmalarına neden olur. Bu noktada devreye MADRL yapıları girer. Bu yapılarda her araç bir ajan gibi davranır; kendi gözlemlerine ve geçmiş deneyimlerine dayanarak politikalarını öğrenir. Aynı zamanda sistem genelinde bir kolektif öğrenme ortamı oluşur. Araçlar

yalnızca bireysel ödüllerle değil, ağıın genel başarımıyla yönlendirilir. Böylece hem bireysel optimizasyon hem de küresel koordinasyon hedeflenebilir.

MADRL'nin en başarılı örneklerinden biri olan MADDPG algoritması, sürekli eylem alanlarında çalışan ajanların merkezi eleştirmen (central critic) yapısı ile eğitilmesini sağlar. Her ajan (araç), kendi yerel gözlemlerinden hareketle aksiyon üretirken; ortak bir değerlendirme yapısı sistemin genel başarısını kontrol eder. Bu yapı sayesinde, kaynak paylaşımı, kanal tahsisi ve iletim stratejileri gibi kararlar hem bağlama duyarlı hem de etkileşimli biçimde optimize edilir (Peng ve ark., 2020).

Tüm bu yapılar düşünüldüğünde, klasik yöntemlerden dinamik sistemlere ve oradan da öğrenen ajanlara geçiş yalnızca performans iyileştirmesi değil; aynı zamanda sistemin zekâ düzeyinin artması, karar kalitesinin gelişmesi ve sürdürülebilirliğinin sağlanması anlamına gelmektedir. Bu dönüşüm, V2X sistemlerini yalnızca veri ileten yapılar olmaktan çıkarıp, karar alabilen, adaptif, optimize edici sistemler hâline dönüştürmektedir. Bu da gelecekteki otonom ulaşım altyapılarının yalnızca fiziksel değil, aynı zamanda bilişsel olarak da esnek olmasını gerektirir.

### 3.5. Derin Takviyeli Öğrenme (DRL)

AI sistemlerinin karar verme süreçlerinde çevresel faktörlere tepki verebilmesi ve bu tepkileri deneyim üzerinden zamanla geliştirebilmesi, özellikle dinamik ve belirsiz ortamlarda büyük bir avantaj sağlamaktadır. Bu bağlamda, RL, ajan adı verilen öğrenen bir yapının, bir çevre ile etkileşim kurarak aldığı aksiyonlar sonucunda ödül veya ceza geri bildirimleri aracılığıyla davranış politikalarını öğrenmesini temel alan bir yöntemdir. RL, çevre durumu (state), eylem (action), ödül (reward) ve politika (policy) gibi temel bileşenlerden oluşur. Her aksiyonun sistemde bir sonucu olur ve bu sonucun olumlu ya da olumsuz geri bildirimi, ajan tarafından sonraki adımlar için öğrenme malzemesi hâline getirilir. Ancak RL algoritmaları yalnızca sınırlı sayıda durum ve eylem içeren küçük ölçekli ortamlarda etkin çalışabilir. Sürekli durum-aksiyon uzayına sahip, çok değişkenli, yoğun veri akışlı ve bağlamsal bilginin önemli olduğu sistemlerde, klasik RL algoritmalarının durum-aksiyon değerlerini depolaması ve güncellemesi mümkün değildir. Çünkü bu tür sistemlerde olası durumlar çok geniştir ve sistem davranışları kesin çizgilerle tanımlanamaz.

Bu zorluk, RL sistemlerinin DL algoritmaları ile birleştirilmesini zorunlu kılmıştır. DNN, yüksek boyutlu veri setlerinde örüntü tanıma, genelleme ve temsili öğrenme konularında oldukça başarılıdır. RL'in bu güçlü temsil kapasitesi ile birleşmesi sonucu ortaya çıkan DRL algoritmaları, geniş ve sürekli uzaylara sahip ortamlarda da etkili politikalar öğrenebilen sistemler geliştirmeye imkân tanır.

DRL'de klasik Q-değeri tabloları yerine bir sinir ağı kullanılarak durum-eylem değerleri tahmin edilir. Bu ağ, örnekleme tabanlı yöntemlerle (örneğin deneyim tekrar havuzu) ve hedef ağırlıklarla eğitilir. Ajan, sistemdeki bir durumu gözlemledikten sonra mevcut politika ağı üzerinden bir eylem önerisi alır, bu eylemin sonucunda çevre bir yeni durum ve ödül üretir. Ajan bu geri bildirimle kendi ağırlıklarını günceller. Sürekli olarak tekrarlanan bu döngü sayesinde ajan, giderek daha iyi bir politika öğrenir (Luong ve ark., 2019). DRL algoritmaları, klasik RL'nin yapılamadığı ortamlarda, hem öğrenme hem de genelleme kapasitesini artırarak, çok boyutlu karar problemlerine etkili çözümler sunar.

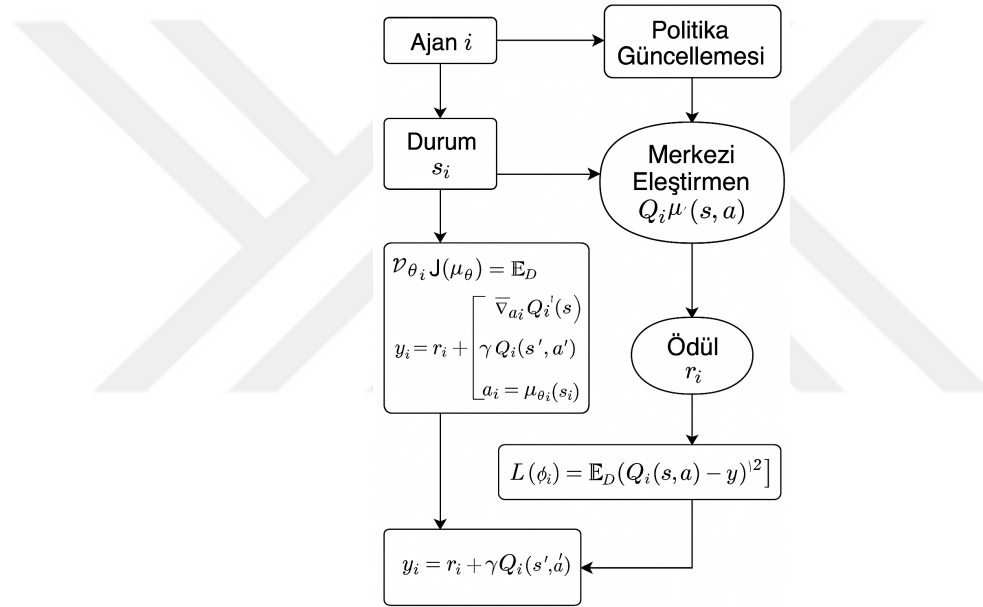
### 3.5.1. MADRL ve MADDPG Yaklaşımı

MADRL yaklaşımları, her bir ajanın kendi başına öğrenme sürecini sürdürmesini sağlarken, aynı zamanda bu ajanlar arasında doğrudan ya da dolaylı bir koordinasyon ve denge kurulmasına olanak tanır. MADRL sistemlerinde ajanlar ya bağımsız olarak öğrenir ya da kolektif bilgi paylaşımı yoluyla ortak hedeflere hizmet eden davranışlar geliştirir.

MADRL'in temel yaklaşımlarından biri olan MADDPG algoritması, özellikle sürekli aksiyon uzaylarında çalışan ajanlar için tasarlanmıştır. MADDPG, klasik DDPG algoritmasının çok etmenli bir varyantıdır ve aktör-eleştirmen (actor-critic) yapısına dayanır. Her bir ajan bir "aktör" olarak kendi politika ağını kullanarak aksiyon üretir. Bu aktörün aksiyonları, sistemdeki bir "merkezî eleştirmen (centralized critic)" tarafından değerlendirilir. Eleştirmen, sistemdeki tüm ajanların durumlarını ve eylemlerini göz önüne alarak küresel bir değerlendirme yapar. Böylece her bir ajan yalnızca kendi çıkarını maksimize etmeye değil, sistemsel başarıya katkıda bulunmaya yönelik politikalar geliştirir. Bu yapı özellikle V2X sistemlerinde avantaj sağlar. Çünkü burada her araç (ajan), çevresel bilgileri kısmen gözlemleyebilir. MADDPG'nin merkezî eleştirmen yapısı, bu eksik gözlemleri tamamlamakta ve sistemin genel

performansına göre ödül yapılarını güncelleyerek bireysel kararları küresel hedefle hizalayabilmektedir.

Şekil 3.3 üzerinde sunulan diyagram, MADDPG algoritmasının ajan bazlı öğrenme döngüsünü göstermektedir. Bu yapıda, her bir ajan  $i$ , ortamdan algıladığı durum  $s_i$ 'yi temel alarak politika fonksiyonu ile bir aksiyon üretmektedir. Ajanın eylemi sonucunda ortamdan aldığı ödül hem merkezi eleştirmen hem de politika güncelleme süreci için kullanılmaktadır. Eleştirmen, tüm ajanların durum ve eylem bilgilerini dikkate alarak Q-değeri üretir; bu değer, hem aktörün gradyan güncellemesi hem de kayıp fonksiyonunun hesaplanmasında kullanılmaktadır. Böylece ajan, bireysel aksiyonunu sistemin ortak performansı doğrultusunda optimize eder.



Şekil 3.3. MADDPG algoritması

Merkezî eleştirmen, her bir ajan için daha kararlı bir öğrenme süreci sağlar. Ortamda diğer ajanların politikaları değişse bile, bu bilgiye erişimi olan eleştirmen genel sistem dinamiğini dengeleyebilir. Ajanlar tam durumu gözlemleyemese bile, merkezî eleştirmen tüm sistem durumunu değerlendirerek öğrenmeyi destekler. V2X gibi sistemlerde iletim gücü, zamanlama, yön seçimi gibi kararlar sürekli değerler alır. MADDPG, bu tür ortamlarda Q-değer tabanlı yaklaşımlardan daha uygundur. V2X sistemlerinin yapısı gereği, her aracın çevresel algılaması farklıdır; fakat hepsi aynı hedef doğrultusunda çalışır. MADDPG sayesinde bu çok ajanlı sistemlerde bireysel çıkarlar ile sistemsel performans dengelenir. Özellikle yüksek bağlantı yoğunluğu,

değişken kanal koşulları ve farklı veri önceliklerinin olduğu ortamlarda MADDPG, klasik DRL ve sabit stratejilere kıyasla çok daha etkin sonuçlar verir (Parvini ve ark., 2023).



## 4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu tez çalışmasında, HAPS destekli bir V2X iletişim senaryosu ele alınmakta; AoI farkında bir kaynak yönetimi politikası önerilmektedir. Bu bölümde; gerçekleştirilen çalışmaların matematiksel temelleri, ele alınan sistem modeli, kullanılan kanal modeli, bilgi güncelliği ölçümleri, optimizasyon problem tanımı ve kullanılan MADDPG algoritması detaylı şekilde sunulmaktadır. Her bir alt bölüm, sistemin işleyişine dair temel bileşenleri açıklamakta ve simülasyon ortamında yürütülen deneysel çalışmaların tekrar üretilebilirliğini sağlayacak düzeyde teknik bilgi içermektedir. Böylece, geliştirilen yaklaşımın hem teorik dayanakları hem de uygulama adımları bütüncül bir bakış açısıyla ortaya konmaktadır.

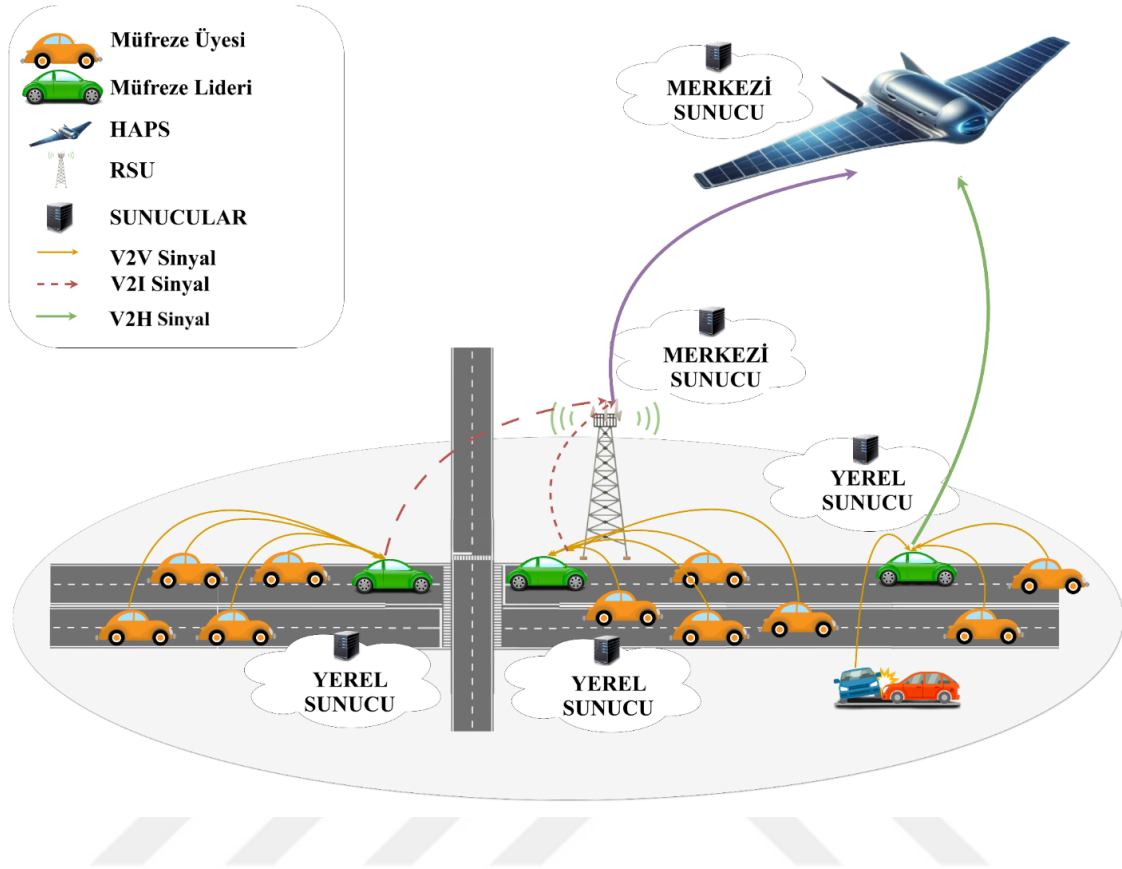
### 4.1. Sistem Mimarisi

Bu çalışmada önerilen sistem mimarisi, V2X ortamında karar destekli kaynak yönetimini hedefleyen, çok katmanlı bir haberleşme yapısına sahiptir. Mimarinin merkezinde, karasal ve karasal olmayan ağ unsurlarının birlikte çalıştığı bir iletişim altyapısı yer almakta; bu yapı üzerinden AoI başta olmak üzere çeşitli performans metriklerini optimize eden bir AI tabanlı öğrenme algoritması çalıştırılmaktadır.

Şekil 4.1'de görüldüğü gibi, ele alınan senaryo, iki boyutlu düzlemde hareket eden çok sayıda otonom araçtan oluşmaktadır. Bu araçlar, belirli bir hızla hareket etmekte ve çevrelerinden topladıkları bağlamsal verilere göre kendi iletişim kararlarını almaktadır. Araçlar hem V2V, hem karayolu üzerindeki RSU aracılığıyla V2I, hem de sistemin merkezinde sabit konumda yer alan HAPS üzerinden V2H iletişim kurabilmektedir. Bu üç katmanlı haberleşme altyapısı, sistemin yalnızca belirli coğrafi alanlarda değil, kapsama dışı bölgelerde de bilgi alışverişini sürdürebilmesini garanti altına almaktadır.

Sistemde yer alan her araç bir “ajan” olarak modellenmiştir. Bu ajanlar, çevrelerindeki bağlantı durumu, kanal kalitesi, AoI, trafik yoğunluğu gibi çok sayıda değişkeni gözlemleyebilmekte; bu gözlemler doğrultusunda uygun haberleşme yolunu, veri iletim zamanlamasını ve enerji kullanımını optimize edecek şekilde aksiyon alabilmektedir. Bu karar süreci sabit kurallarla değil, RL temelli çok etmenli bir mimari ile desteklenmektedir. Bu sayede ajanlar yalnızca geçmiş deneyimlerinden öğrenmekle

kalmamakta; aynı zamanda diğer ajanların davranışlarını da göz önünde bulundurarak kolektif düzeyde bilgi güncelliğini korumaya çalışmaktadır.



Şekil 4.1. Sistem senaryosu

HAPS, sabit bir irtifada konumlandırılmış olup geniş bir kapsama alanına sahiptir. Tüm araçlar ile LoS bağlantısı kurabilmekte, böylece özellikle karasal altyapının yetersiz kaldığı durumlarda veya trafik yoğunluğunun RSU kapsama sınırlarını aştığı senaryolarda veri iletimini kesintisiz hâle getirmektedir. HAPS'in sistem üzerindeki etkisi yalnızca bağlantı sürekliliği ile sınırlı değildir; aynı zamanda sinyal zayıflama değerlerinin daha stabil olması sayesinde, öğrenen ajanların karar alma süreçlerinde daha az belirsizlik oluşmakta ve öğrenme eğrileri daha istikrarlı seyretilmektedir.

Sistem mimarisi, her zaman adımında (discrete timestep) güncellenen bir yapıya sahiptir. Bu adımlar boyunca araçların konumları değişmekte, bağlantı durumları güncellenmekte, bilgi güncelliği hesaplanmakta ve her ajan kendi gözlemine dayalı olarak bir aksiyon seçmektedir. Bu yapının doğrudan sonucu olarak ortaya çıkan veri akışı, çok etmenli bir pekiştirmeli öğrenme süreciyle yönetilmekte; sistem politikaları zaman içerisinde çevreye uyum sağlayarak gelişmektedir.

Bu mimari yapının matematiksel temelleri, kanal modeli, AoI hesaplamaları ve optimizasyon probleminin tanımlanmasıyla birlikte bir sonraki alt bölümlerde ayrıntılı biçimde ele alınacaktır.

## 4.2. Kanal Modeli

Geliştirilen HAPS destekli V2X sisteminde, bilgi iletimi farklı fiziksel katman özelliklerine sahip üç temel bağlantı yolu üzerinden gerçekleştirilmektedir: V2V, V2I ve V2H. Her bir bağlantı türü, kendine özgü zayıflama (path loss), gecikme ve kapsama özelliklerine sahiptir. Bu nedenle, haberleşme ortamının fiziksel katman modellemesi, bağlantı seçim stratejilerinin doğruluğu ve sistem performans ölçümlerinin geçerliliği açısından kritik öneme sahiptir.

Bu çalışmada, sinyal zayıflaması ve kanal kalitesi hesaplamaları için logaritmik mesafeye dayalı zayıflama modeli (log-distance path loss model) kullanılmıştır. Bu model, hem karasal hem de HAPS bağlantıları için parametreleri özelleştirilmiş biçimde uygulanmıştır.

### 4.2.1. Karasal Bağlantı Kanalları (V2V ve V2I)

V2V ve V2I haberleşmesi, sistemin karasal bağlantı katmanını oluşturmaktadır. Bu tür bağlantılar genellikle daha düşük mesafeli ancak yüksek gecikme hassasiyetine sahip olup, özellikle düşük trafik yoğunluğuna sahip bölgelerde ve RSU kapsam alanında yer alan araçlar arasında tercih edilmektedir. Her iki bağlantı türü için benzer fiziksel kanal özellikleri geçerli olduğundan, aynı matematiksel model altında birlikte ele alınmaktadır. Karasal kanal modeli, sinyal gücünün iletim mesafesi arttıkça azaldığı gerçeğine dayanan klasik logaritmik zayıflama modeli ile temsil edilmektedir. Bu modelde, iki haberleşme noktası arasındaki zayıflama miktarı, logaritmik olarak mesafeye bağlı şekilde Eşitlik 4.1'de ki gibi modellenir (Alfattani ve ark., 2021):

$$PL_{V2X}(d) = PL_0 + 10n \log_{10}(d/d_0) \quad (4.1)$$

Burada  $PL_{V2X}(d)$ ,  $d$  mesafesinde ölçülen toplam zayıflama miktarı (dB cinsinden),  $PL_0$  referans mesafede ( $d_0$ ) ölçülen başlangıç zayıflaması (dB),  $n$  ortamın yayılım katsayısı (örneğin açık alan için  $n = 2$ , kentsel alanlar için  $n = 3 \sim 4$ ),  $d$  iki haberleşme noktası arasındaki doğrusal mesafe (metre),  $d_0$  ise referans mesafeyi ifade eder

Simülasyon ortamında, Eşitlik 4.2 de gösterildiği gibi, her araç  $i$  zamanın  $t$  anında iki boyutlu düzlemde bir konuma sahiptir.

$$p_i(t) = (x_i(t), y_i(t)) \quad (4.2)$$

İki araç  $i$  ve  $j$  arasındaki doğrusal mesafe Eşitlik 4.3'te gösterildiği gibi hesaplanır

$$d_{ij}(t) = (p_i(t) - p_j(t)) = \sqrt{(x_i(t) - x_j(t))^2 + (y_i(t) - y_j(t))^2} \quad (4.3)$$

Bu mesafe değeri, Eşitlik 4.1'de  $d$ 'nin yerine yazılarak, bağlantının zayıflama düzeyi elde edilir. Aynı işlem, araç ile RSU arasındaki bağlantılar için de geçerlidir.

#### 4.2.2. HAPS Bağlantısı (V2H)

Sistem mimarisinde yer alan HAPS, yüksek irtifa konumlandırılması sayesinde araçlar ile LoS kurabilen, geniş kapsama alanına sahip bir iletişim birimi olarak modellenmiştir. HAPS ile araçlar arasındaki haberleşme, özellikle karasal altyapının yetersiz olduğu veya trafik yoğunluğunun yüksek olduğu durumlarda kritik önem taşımaktadır. Bu nedenle, HAPS kanal modeli yalnızca mesafe tabanlı değil; aynı zamanda görüş hattı olasılığı, frekans bağımlı yayılım ve ortam kaynaklı bozulmalar gibi çok sayıda fiziksel parametreyi içerecek biçimde modellenmiştir.

HAPS bağlantısında kanal kazancı, Ricean sönümlenme kabulüyle Eşitlik 4.4'te gösterildiği şekilde modellenir:

$$h_{j,s}^t[k] = 10^{-PL/20} \left( \sqrt{p_L} a_j^t + \sqrt{p_N} h_j^t[k] \right) \quad (4.4)$$

Burada,  $h_{j,s}^t[k]$  HAPS ile araç  $j$  arasındaki  $k$ . alt kanalın kanal kazancı,  $p_L$  LoS olasılığı,  $a_j^t = \exp\left(j \frac{2\pi d_{ij}}{\lambda}\right)$  deterministik LoS bileşeni,  $h_j^t[k]$  LoS olmayan (non-LoS, N-LoS) sönümlenme bileşeni (kompleks Gauss değişkeni),  $\lambda$  ise sinyal dalga boyunu ifade eder.

LoS olasılığı, HAPS'in belirli bir yükselme açısı  $v$  ile araca göre konumlandığı varsayılarak aşağıdaki Eşitlik 4.5'te tanımlanmıştır:

$$p_L = \frac{b_1 v}{b_2 + b_3} \quad (4.5)$$

Burada,  $b_1, b_2, b_3$  çevresel parametrelere göre belirlenen sabit katsayılardır,  $v$ : HAPS ile araç arasındaki yükselme açısıdır. Yükselme açısı, HAPS yüksekliği  $H$  ve yatay mesafe  $d_H$  dikkate alınarak Eşitlik 4.6'da gösterildiği gibi hesaplanır:

$$v = \arctan\left(\frac{H}{d_H}\right) \quad (4.6)$$

Toplam  $PL$  değeri, LoS ve NLoS olasılıklarının ağırlıklı ortalamasıyla Eşitlik 4.7'de gösterildiği gibi hesaplanır:

$$PL = p_L \cdot PL_{LoS} + (1 - p_L) \cdot PL_{NLoS} \quad (4.7)$$

Her bir bileşen  $PL_z$ ,  $z \in \{LoS, NLoS\}$  olmak üzere, Eşitlik 4.8'de ki gibi ayrıştırılır:

$$PL_z = FS + C_z + \chi_z \quad (4.8)$$

Burada,  $FS$  serbest alan zayıflaması (free space path loss),  $C_z$  ortam türüne bağlı sabit katsayı (LoS/NLoS için farklı),  $\chi_z$  gölgelenme (shadowing) etkisini temsil eden Gauss dağılımlı rassal değişkeni ifade eder.  $FS$ , frekansa ve mesafeye bağlı olarak Eşitlik 4.9'da ki gibi tanımlanır:

$$FS = 32.45 + 20 \log_{10}(f) + 20 \log_{10}(d) \quad (4.9)$$

Burada,  $f$ : taşıyıcı frekans (MHz),  $d$ : HAPS ile araç arasındaki doğrusal mesafe (km) ifade eder. Doğrusal mesafe  $d$ , yerin eğriliği de hesaba katılarak Eşitlik 4.10'da ki gibi hesaplanır:

$$d = \sqrt{R_E^2 \sin^2(\theta) + H^2 + 2HR_E - R_E \sin(\theta)} \quad (4.10)$$

Burada,  $R_E$  yer yarıçapı,  $\theta$  araçla HAPS arasındaki açısal irtifayı ifade eder.

Bu hesaplama, HAPS'in konumlandığı irtifaya ve yer eğriliğine bağlı olarak daha doğru bir sinyal yayılım modeli sağlar.

Bu model sayesinde HAPS bağlantısı, yalnızca statik bir "yüksekten sinyal alanı" değil, çevresel değişkenlere göre şekillenen ve LoS olasılığıyla sürekli güncellenen bir kanal olarak simüle edilmektedir. Bu da öğrenen ajanların, HAPS bağlantısını sadece acil durum alternatifi değil, bağlamsal olarak tercih edilebilir bir seçenek olarak değerlendirmesine olanak tanır.

#### 4.2.3. Bağlantı Eşiği ve Kanal Kullanılabilirliği

Haberleşme kalitesinin sürdürülebilirliği açısından, sistemde yalnızca belirli bir sinyal zayıflama eşiğini karşılayan bağlantıların aktif olması kabul edilmiştir. Bu yaklaşım, yalnızca fiziksel bir gereklilik değil, aynı zamanda sistemin performans optimizasyonunu sağlayan önemli bir kontrol mekanizmasıdır. Gerek V2V/V2I gibi karasal bağlantılar, gerekse V2H gibi HAPS tabanlı bağlantılar için sinyal gücü, her bir zaman adımında hesaplanan toplam yol kaybına göre değerlendirilmekte; bu değer sistem tarafından belirlenen bir eşik değeriyle karşılaştırılmaktadır.

Bir bağlantının kullanılabilir (aktif) kabul edilmesi için Eşitlik 4.11'deki koşul sağlanmalıdır:

$$PL(d) \leq PL_{th} \quad (4.11)$$

Burada,  $PL(d)$  bağlantı için hesaplanan toplam zayıflama değeri,  $PL_{th}$  ise kanal erişim eşiğini ifade etmekte ve bağlantının kabul edilebilir sinyal kalitesine sahip olup olmadığını belirleyen sistem parametresi olarak değerlendirilmektedir.

Bu koşul yalnızca fiziksel olarak mümkün olan bağlantıları seçmekle kalmaz, aynı zamanda bağlantı topolojisinin zamanla değişmesini de sağlar. Araçlar hareket ettikçe aralarındaki mesafeler, dolayısıyla zayıflama değerleri de değişmekte ve bu dinamik yapı sayesinde bağlantı grafiği sürekli güncellenmektedir.

### 4.3. AoI Modeli

Zaman duyarlılığı yüksek iletişim ortamlarında, yalnızca paket iletim süresi değil, alınan bilginin ne kadar güncel olduğu da sistem performansını doğrudan etkileyen bir faktördür. Özellikle dinamik ağ yapılarında, bir aracın diğerlerinden aldığı bilgilerin güncelliği, kontrol kararlarının doğruluğu ve güvenliği açısından kritik rol oynar. Bu nedenle, klasik gecikme, bant genişliği ve paket kaybı gibi ölçütlerin yanında, bu çalışmada AoI metriği temel bir performans göstergesi olarak kullanılmıştır.

Bu çalışmada her ajan, kendisine ait verilerin ve çevresinden edindiği bilgilerin güncelliğini takip eder. Her ajan için AoI, ayrık zaman adımlarıyla ilerleyen sistemde anlık olarak güncellenir. Anlık AoI değeri, en son alınan verinin üretim zamanına göre tanımlanır ve Eşitlik 4.12’de ki gibi hesaplanır:

$$\Delta_i(t) = t - u_i(t) \quad (4.12)$$

Burada,  $\Delta_i(t)$  ajan  $i$ 'nin  $t$  anındaki anlık AoI değerini ifade ederken  $u_i(t)$  ajan  $i$ 'nin aldığı en son verinin üretildiği zaman anını temsil eder.

Bu denklem doğrultusunda, eğer ajan  $i$ ,  $t$  anında yeni bir güncelleme almazsa,  $u_i(t) = u_i(t-1)$  kalır ve  $\Delta_i(t) = \Delta_i(t-1) + 1$  olur. Eğer ajan  $i$ ,  $t$  anında güncel bir veri paketi alırsa,  $u_i(t) = t$  olur ve  $\Delta_i(t) = 0$  'dır.

Bu yapı sayesinde, sistemdeki her ajan, hem kendi bilgisinin hem de iletişim kurduğu diğer birimlerin bilgi tazeliğini izleyebilmekte ve öğrenme sürecine bağlamsal bilgi güncelliği temelinde katkıda bulunmaktadır.

Anlık AoI değeri, zaman ilerledikçe artan ve güncelleme alındığında sıfırlanan yapısıyla tipik olarak testere dişi (sawtooth) biçiminde bir zamana bağlı fonksiyon üretir. Bu eğilim, veri trafiğinin sıklığını, düzenliliğini ve gecikmelerini doğrudan görselleştirme imkânı da sunmaktadır.

Anlık AoI değeri, bir aracın belirli bir anda elinde bulundurduğu bilginin güncelliğini ölçmek açısından önemli bir gösterge olmakla birlikte, sistemin genel başarımını değerlendirmek için tek başına yeterli değildir. Bu nedenle, birden fazla zaman adımı boyunca toplanan AoI değerlerinden türetilen özet istatistikler olan ortalama AoI ve maksimum AoI, sistemin bilgi tazeliği üzerindeki bütünsel etkisini analiz etmek amacıyla kullanılmaktadır.

Ortalama AoI, bir ajan veya sistemin belirli bir zaman aralığında ne kadar güncel bilgiyle çalıştığının ortalamasıdır. Simülasyon süresi boyunca ajan  $i$ 'nin AoI değerlerinin ortalaması Eşitlik 4.13 'te ki gibi tanımlanır.

$$\bar{\Delta}_i = \frac{1}{K} \sum_{k=1}^K \Delta_i(k) \quad (4.13)$$

Bu eşitlikte,  $\bar{\Delta}_i$  ajan  $i$ 'nin simülasyon süresince gözlemlenen ortalama AoI değeri,  $K$  toplam zaman adımı sayısı,  $\Delta_i(k)$  ajan  $i$ 'nin  $k$ . adımda sahip olduğu anlık AoI değeri.

Bu metrik, her ajan için ayrı ayrı hesaplanarak sistemin bireysel bilgi güncelliği performansını değerlendirmeye imkân tanır.

Toplam sistem performansını değerlendirmek amacıyla,  $N$  toplam ajan sayısını göstermek üzere tüm ajanların ortalama AoI değerleri üzerinden bir genel sistem ortalaması Eşitlik 4.14'te ki gibi hesaplanır.

$$\bar{\Delta}_{system} = \frac{1}{N} \sum_{i=1}^N \bar{\Delta}_i \quad (4.14)$$

Bu ifade, sistemin tamamı için bilgi güncelliği açısından ortalama bir başarı ölçütü sunar. Özellikle farklı iletişim stratejilerinin karşılaştırılmasında ve öğrenme algoritmalarının etkinliğinin analizinde bu metrik temel gösterge olarak kullanılır.

Maksimum AoI, sistemin herhangi bir ajanı için oluşabilecek en yüksek AoI değerini ifade eder. Bu değer, özellikle güvenlik kritik sistemlerde veya gecikmeye karşı toleranssız uygulamalarda sistemin en kötü hâlini yansıttığı için Eşitlik 4.15 önemlidir.

$$\Delta_{max} = \max_{i,t} \Delta_i(t) \quad (4.15)$$

Bu hesaplama, simülasyon boyunca tüm ajanlar ve tüm zaman adımları için Eşitlik 4.16'da gösterdiği gibi yapılır:

$$\Delta_{max} = \max \Delta_i(t); i=1, \dots, N; t \in [0, T] \quad (4.16)$$

Maksimum AoI değeri, özellikle öğrenme sürecinde sınırlandırılması gereken bir risk faktörü olarak değerlendirilir. Öğrenme algoritmalarının sadece ortalama AoI'yı değil, aynı zamanda bu en kötü senaryoları da minimize etmeyi hedeflemesi, sistemin daha kararlı ve güvenli çalışmasını sağlar.

Bilgi güncelliği, yalnızca sistemin analizinde kullanılan pasif bir metrik değil, aynı zamanda geliştirilen öğrenme algoritmasının doğrudan yönlendirdiği ve iyileştirmeyi hedeflediği bir performans değişkenidir. Bu bağlamda, AoI yalnızca gözlemlenen bir çıktı değil; ajanların eylem seçimlerini etkileyen bir öğrenme hedefi hâline getirilmiştir.

Bu çalışmada, çok etmenli pekiştirmeli öğrenme çerçevesinde kullanılan MADDPG algoritması, ajanların çevrelerinden elde ettikleri gözlemler ve aldıkları ödüller aracılığıyla zamanla daha verimli karar politikaları geliştirmesini amaçlamaktadır. AoI değeri, ajanların aldığı ödül fonksiyonuna doğrudan entegre edilerek, her ajanın bilgi güncelliğini mümkün olan en düşük seviyede tutmaya yönelik davranışlar öğrenmesi sağlanmaktadır. Eşitlik 4.17'de AoI'nın ödül fonksiyonuna nasıl entegre edildiği gösterilmiştir.

$$r_i(t) = -\omega_{\Delta} \cdot \Delta_i(t) + \phi_i(t) \quad (4.17)$$

Burada,  $r_i(t)$  ajan  $i$ 'nin  $t$  anındaki toplam ödülü,  $\Delta_i(t)$  ajan  $i$ 'nin AoI değerini,  $\omega_{\Delta}$  AoI'ya verilen önem katsayısını (pozitif bir skaler),  $\phi_i(t)$  diğer pozitif katkı sağlayan ödül bileşenleri (örneğin enerji verimliliği, bağlantı başarı oranı vb.) temsil etmektedir.

Bu yapı sayesinde, AoI yükseldikçe ödül değeri düşer, ajanlar yalnızca bağlantı kurmakla kalmaz, ne zaman ve hangi bağlantıyla bilgi güncellemesi yapmaları gerektiğini öğrenir. Ajanların eylem uzayı yalnızca sabit seçeneklerden oluşmaz; her zaman adımında değişen kanal kullanılabilirliği, mesafe, bağlantı kalitesi ve geçmiş AoI değerleri gibi bağlamsal bilgilerle şekillenir. Bu durum, ajanların öğrenme süreci boyunca yalnızca kurallara dayalı değil, ortama duyarlı ve duruma özgü politikalar geliştirmelerine olanak tanır. Örneğin; eğer AoI değeri düşükse ve bağlantı kalitesi kötüyse, ajan bağlantı kurmamayı öğrenebilir. Eğer AoI değeri yüksekse, o anda aktif olan en uygun bağlantı kanalını kullanmayı tercih edebilir. Bu öğrenme süreci sonunda ajanlar, hem bilgi güncelliğini yüksek tutan hem de ağ kaynaklarını etkin kullanan bir iletişim stratejisi geliştirir.

AoI'nın öğrenme sürecine entegre edilmesi sayesinde, ajanlar daha düzenli bir şekilde ve optimum zamanlamayla bilgi güncellemesi yapar, gereksiz bağlantılar ve enerji israfı önlenir. Sistem genelinde hem ortalama hem de maksimum AoI değerlerinde anlamlı düşüş sağlanır. Bu yapı, yalnızca kısa vadeli başarıyı değil, uzun vadede sistemin kararlılığını, uyarlanabilirliğini ve verimliliğini de artırır.

Geliştirilen HAPS destekli V2X haberleşme sisteminde, haberleşme kanallarının sınırlı bant genişliği, değişken kanal koşulları ve eşzamanlı kaynak talepleri gibi kısıtlar altında, her bir aracın ne zaman, hangi iletişim modunu kullanacağına, hangi güç seviyesinde iletim yapacağına ve hangi frekans dilimini seçeceğine karar vermesi gerekmektedir. Bu kararlar, hem bireysel ajan performanslarını hem de sistem genelindeki bilgi tazeliği başarımını doğrudan etkilemektedir.

Bu bağlamda, sistemin davranışı matematiksel olarak bir optimizasyon problemi şeklinde tanımlanmıştır. Problem; iletişim mod seçimi, spektrum tahsisi ve güç kontrolünü içeren karar değişkenleri üzerinden, toplam AoI'nın minimize edilmesini hedeflemektedir.

#### **4.3.1. İletişim Kapasitesi**

Geliştirilen HAPS destekli V2X sisteminde modlar arasında seçim, o andaki kanal koşulları, sinyal seviyeleri ve sistem politikalarına bağlı olarak belirlenmektedir. Her mod, kendine özgü bir kanal modeli, sinyal zayıflama yapısı ve parazit etkisi

altında çalışmaktadır. Bu nedenle her iletişim modu için ayrı kapasite tanımları yapılması, sistemin optimizasyon problemine temel teşkil eden kapasite yapılarının doğru ve bütüncül olarak modellenmesini sağlar.

Her mod için kapasite hesaplamaları, Shannon kapasite teorisine dayalı olarak yapılmış olup; kaynak dilimi başına düşen kapasite değeri, iletim gücü, kanal kazancı ve parazit seviyesi gibi fiziksel parametreleri içermektedir. Bu bölümde, üç iletişim modu için kapasite fonksiyonları ayrı ayrı ele alınmakta ve delta fonksiyonları aracılığıyla mod seçim durumu matematiksel olarak modele entegre edilmektedir.

#### 4.3.1.1. V2I Kanal Kapasitesi

V2I modunda, araçlar doğrudan yol kenarına konumlandırılmış sabit RSU istasyonlarıyla haberleşir. Bu iletişim türü, sabit ve güçlü altyapı bağlantısı sağladığı için genellikle yüksek kapasite sunabilir. Ancak erişilebilirliği RSU'nun kapsama alanı ile sınırlıdır. Zaman  $t$ 'de, araç  $j$ 'nin  $k$ . frekans kanalında V2I modunda elde ettiği kapasite Eşitlik 4.18'de verildiği gibi tanımlanır:

$$C_{tj,I}[k] = \log_2 \left( 1 + \frac{\delta(\theta_j) \beta_{j,k}^t p_j^t[k] h_{j,I}^t[k]}{I_{j,I}[k] + \sigma^2} \right) \quad (4.18)$$

Burada,  $\theta_j \in \{0,1,2\}$ : araç  $j$ 'nin o anda çalıştığı iletişim modunu belirtir.  $\theta_j = 0$  olduğunda araç V2I modundadır,  $\delta(\theta_j)$  delta fonksiyonu,  $\beta_{j,k}^t \in \{0,1\}$  araç  $j$ 'nin  $k$ . frekans kanalını zaman  $t$ 'de kullanıp kullanmadığını belirten ikili değişken,  $p_j^t[k]$  araç  $j$ 'nin o frekans kanalına ayırdığı iletim gücü,  $h_{j,I}^t[k]$  araç ile RSU arasındaki kanal kazancı,  $\sigma^2$ : beyaz Gauss gürültü gücüdür.  $I_{j,I}[k]$  aynı frekans kanalını kullanan diğer araçlardan kaynaklanan paraziti ifade eder ve Eşitlik 4.19'da ki gibi tanımlanır:

$$I_{j,I}[k] = \sum_{j' \neq j} \beta_{j',k}^t p_{j'}^t[k] h_{j',I}^t[k] \quad (4.19)$$

Bu yapı, araçların aynı altyapı kanalını paylaşmaları durumunda oluşan parazit etkisini dikkate alarak, daha gerçekçi bir kapasite tahmini yapılmasını sağlar.

#### 4.3.1.2. V2V Kanal Kapasitesi

V2V iletişim modunda, araçlar doğrudan birbirleriyle haberleşerek bilgi paylaşımını gerçekleştirir. Bu mod, altyapı birimlerinin bulunmadığı ortamlarda ya da RSU bağlantısının zayıf olduğu durumlarda kritik öneme sahiptir. Özellikle araç filoları, güvenli geçiş koordinasyonu veya lokal trafik bilgisi paylaşımı gibi durumlar için V2V bağlantıları vazgeçilmezdir.

Ancak, bu bağlantı türü yüksek hareketlilik, düzensiz kanal koşulları ve yoğun parazit etkisi altında çalıştığından, kapasite tahmini dikkatle yapılmalıdır. V2V modundaki iletişim kapasitesi, yine Shannon kapasitesi temel alınarak Eşitlik 4.20'de gösterildiği şekilde tanımlanır:

$$C_{j,V}^t[k] = \log_2 \left( 1 + \frac{\delta(\theta_j - 1) \beta_{j,k}^t p_j^t[k] h_{j,V}^t[k]}{I_{j,V}[k] + \sigma^2} \right) \quad (4.20)$$

Burada,  $\theta_j = 1$  olduğunda araç V2V modundadır.  $I_{j,V}[k]$  aynı frekans kanalını kullanan diğer araçlardan kaynaklanan parazit olup Eşitlik 4.21'deki gibi tanımlanır:

$$I_{j,V}[k] = \sum_{j' \neq j} \beta_{j',k}^t p_{j'}^t[k] h_{j',V}^t[k] \quad (4.21)$$

V2V kapasitesi, bağlantı mesafesi kısa olduğunda yüksek olabilir; ancak aynı anda çok sayıda araç V2V iletim yapıyorsa, parazit seviyesi ciddi biçimde artabilir ve bu da kapasiteyi düşürür. Bu nedenle bu mod, kanal paylaşımına en duyarlı olan yapı olarak değerlendirilir.

#### 4.3.1.3. V2H Kanal Kapasitesi

HAPS, geleneksel yer tabanlı altyapılardan farklı olarak geniş kapsama alanı sağlayan ve LoS ihtimalinin yüksek olduğu hava tabanlı iletişim platformlarıdır. Bu sistemler, özellikle RSU dışı bölgelerde veya yoğun V2V/V2I trafiğinin aksamaya uğradığı anlarda yedek bağlantı ve kapsama devamlılığı açısından stratejik bir rol üstlenir.

V2H iletişimde kanal yapısı genellikle Ricean fading etkisine açıktır ve yükselme açısı, atmosferik koşullar gibi faktörlerden etkilenebilir. Sistemde, bu tür bağlantının kapasitesi Eşitlik 4.22’de modellenmiştir:

$$C_{j,H}^t[k] = \log_2 \left( 1 + \frac{\delta(\theta_j - 2) \beta_{j,k}^t p_j^t[k] h_{j,H}^t[k]}{I_{j,H}[k] + \sigma^2} \right) \quad (4.22)$$

Burada,  $\theta_j = 2$  olduğunda araç V2H modundadır.  $I_{j,H}[k]$  aynı frekans kanalını kullanan diğer araçlardan kaynaklanan paraziti temsil eder ve Eşitlik 4.23’te gösterildiği gibi tanımlanır:

$$I_{j,H}[k] = \sum_{j' \neq j} \beta_{j',k}^t p_{j'}^t[k] h_{j',H}^t[k] \quad (4.23)$$

Eşitlik 4.23’te, ajan  $j$  dışındaki tüm ajanların aynı frekans dilimini kullanarak HAPS üzerinden yaptıkları iletimlerin yarattığı girişimi yansıtır.

Bu bağlantı modunun avantajı, çok geniş bir coğrafi alanı kapsayabilmesi ve LoS olasılığının yüksek olmasıdır. Ancak, sınırlı spektrum kapasitesi ve enerji maliyeti açısından daha seçici kullanılması gereken bir moddur. Bu nedenle sistemin öğrenme sürecinde, V2H modunun yalnızca gerekli durumlarda devreye alınması amaçlanır.

### 4.3.2. AoI Güncelleme Kuralları

Bu çalışmada önerilen HAPS destekli V2X haberleşme mimarisi kapsamında, her bir aracın güncel bilgiye erişim süresi, yani AoI değeri, iletişim moduna ve haberleşme başarımına bağlı olarak zamanla dinamik biçimde güncellenmektedir. AoI güncellemeleri, yalnızca mod seçimiyle değil, aynı zamanda kanal kalitesi ve iletim kapasitesinin belirli eşikleri karşılayıp karşılamadığıyla da doğrudan ilişkilidir. Her ajan  $j$  için tanımlanan AoI değeri, Eşitlik 4.24’te gösterilen üç koşullu ifade ile güncellenir:

$$A_j^{t+1} = \begin{cases} \Delta t, & \theta_j=0 \text{ ve } \beta_{j,I}^t C_{j,I}^t[k] \geq C_{j,I}^{min} \\ \Delta t, & \theta_j=2 \text{ ve } \beta_{j,S}^t C_{j,S}^t[k] \geq C_{j,S}^{min} \\ A_j^t + \Delta t, & \text{diğer} \end{cases} \quad (4.24)$$

Burada  $\Delta t$ , bir zaman adımının süresidir ve sabittir.  $\theta_j \in \{0,1,2\}$  ajan  $j$ 'nin haberleşme modu seçimini ifade eder.  $C_{j,I}^t[k]$  ve  $C_{j,S}^t[k]$  ajan  $j$ 'nin RSU ve HAPS ile olan iletişim kapasitesidir.

Eşitlik 4.24'te tanımlanan kurala göre, eğer ajan, V2I modunda iletişim kurmaktaysa ve bu iletişim için tahsis edilen kapasite belirlenen alt sınırdan büyükse, veri güncellenmiş kabul edilir ve AoI sıfırlanarak  $\Delta t$ 'ye eşitlenir. Benzer şekilde, V2H modu için de minimum kapasite eşiği karşılandığında AoI sıfırlanır. Ancak bu iki koşul sağlanmadığında (örneğin V2V kullanıldığında ya da yeterli kapasite sağlanamadığında), ajan yeni veri alamadığı için mevcut AoI değeri zaman adımı kadar artırılır. Bu yapı, hem iletişim modlarının performansını hem de kanal kalitesini doğrudan AoI güncelleme sürecine entegre ederek, sistemin bilgi güncelliğini daha hassas şekilde yönetmesini sağlar. Aynı zamanda bu ifade, optimizasyon probleminde hedef fonksiyonun temel dayanağını oluşturur.

### 4.3.3. Optimizasyon Probleminin Formülizasyonu

Bu tez kapsamında önerilen HAPS destekli V2X haberleşme sistemi, bilgi güncelliğini artırmak, spektrum kullanımını optimize etmek ve enerji verimliliğini sağlamak amacıyla çok kriterli bir optimizasyon yapısı ile modellenmiştir. Sistem, araç PL'lerinin heterojen haberleşme ortamında hem iletişim modunu (V2I, V2V, V2H), hem alt kanal tahsisini hem de iletim gücünü eşzamanlı olarak belirleyerek zamana bağlı bir karar mekanizması yürütür.

Bu karar sürecinin temel amacı, her araç için AoI değerini minimize ederken aynı zamanda kanal kapasitesi gereksinimlerini karşılamak ve toplam iletim gücünü sınırlar dâhilinde tutmaktır. Bu bağlamda önerilen optimizasyon problemi, mevcut kısıtlar çerçevesinde Eşitlik 4.25'te tanımlanmıştır:

$$\min_{\beta, \theta, p} \left\{ \frac{1}{T} \sum_{t=1}^T A_j^t - \Pr \left\{ \sum_{t=1}^T \sum_{k \in K} \min_i \{ C_{j,V}^t[k] \} \Delta t \geq \zeta_j \right\} + \frac{1}{T} \sum_{t=1}^T \sum_{k \in K} p_j^t[k] \right\}$$

$$\begin{aligned} K1: & C_{j,S}^t[k] \geq C_{j,S}^{\min}, \forall j \in P, \forall k \in K, \\ K2: & C_{j,I}^t[k] \geq C_{j,I}^{\min}, \forall j \in P, \forall k \in K, \\ K3: & \beta_{j,i}^t \in \{0,1\}, \theta_j \in \{0,1,2\}, \forall j \in P, \forall i \in I, S \quad (4.25) \\ K4: & \sum_{k \in K} \beta_{j,i}^t \leq 1, \forall j \in P, \forall t \in N \\ K5: & p_j^t[k] \leq p_j^{\max}, \forall j \in P, \forall k \in K \end{aligned}$$

Eşitlik 3.25'te, her bir araç  $j$  için tanımlanan hedef; bilgi güncelliğini temsil eden ortalama AoI değerini minimize etmek, minimum kapasite gereksinimlerini sağlayan iletim olasılığını artırmak ve toplam iletim gücünü azaltmaktır. İlk terim, tüm zaman boyunca ajanların ortalama AoI değerlerinin minimize edilmesini amaçlarken, ikinci terim belirli bir iletim kapasitesi eşiği,  $\zeta_j$ , sağlandığında başarılı veri iletiminin gerçekleşme olasılığını ifade eder. Üçüncü terim ise ajanların frekans kaynaklarına tahsis ettikleri toplam iletim gücünü temsil eder ve enerji verimliliği açısından minimize edilmek istenir. Bu hedef fonksiyon, dört temel kısıt altında çözülmektedir: K1 ve K2, her ajan için seçilen iletişim modunda kapasitenin belirli bir eşik değerin üzerinde olmasını şart koşar. K3, kanal ve mod seçim değişkenlerinin ikili (binary) yapıda olmasını ve yalnızca bir modun seçilmesini garanti eder. K4, her ajan için her zaman adımında yalnızca bir alt kanalın seçilmesine izin verir. K5 ise ajanların her bir kaynak kanalına tahsis edebileceği maksimum iletim gücünü sınırlar. Bu yapı, hem bilgi tazeliğini hem de iletişim kalitesini korurken, sınırlı kaynakların etkili yönetimini sağlayan bütünsel bir kontrol çerçevesi sunar.

#### 4.4. MADDPG Tabanlı Çözüm

Önceki bölümde matematiksel olarak tanımlanan çok amaçlı optimizasyon probleminin doğası, klasik deterministik veya sezgisel algoritmalarla çözümü güçleştiren birçok özellik barındırmaktadır. Bunlar arasında sürekli ve ayrık karar değişkenlerinin bir arada bulunması, ajanlar arası etkileşimlerin olması, iletişim ortamındaki belirsizlikler ve dinamik kanal koşulları öne çıkmaktadır. Bu nedenle, önerilen sistemin etkin biçimde yönetilebilmesi için öğrenmeye dayalı, çevik ve ölçeklenebilir bir yaklaşıma ihtiyaç duyulmaktadır.

Bu hususlar dikkate alınarak, tez çalışmasında MADDPG algoritması tercih edilmiştir. MADDPG, çoklu etkenlerin (ajanların) ortak bir ortamda birbirleriyle etkileşim içinde olduğu sistemlerde, her bir ajanın kendi eylem politikasını öğrenmesini sağlayan, aktör-eleştirmen (actor-critic) tabanlı bir pekiştirmeli öğrenme yöntemidir. Bu algoritma, merkezi bir eğitim süreci (centralized training) ile tüm ajanların bilgilerini kullanarak öğrenme gerçekleştirirken, dağıtık (decentralized) bir şekilde karar alarak sistem içindeki karmaşıklığı azaltır. Bu özellik, özellikle V2X ortamında çok sayıda aracın eş zamanlı olarak haberleşme kaynaklarını kullanmaya çalıştığı senaryolarda büyük avantaj sağlamaktadır.

MADDPG algoritmasının bu sistemdeki uygulaması, her aracın hem kendi durumuna hem de çevresel gözlemlerine dayalı olarak iletişim modunu, alt kanal tahsisini ve iletim gücünü öğrenmesine imkân tanımaktadır. Öğrenilen politikalar sayesinde ajanlar, zamanla AoI değerlerini minimize ederken, spektrum kaynaklarını daha verimli kullanmakta ve toplam iletim gücünü kontrol altında tutmaktadır. Bu bölümde, algoritmanın sistem üzerindeki uygulanma biçimi, ajan başına tanımlanan durum-eylem-ödül uzayları ve öğrenme süreçleri ayrıntılı olarak ele alınacaktır.

Önerilen sistemde her bir ajan , zaman adımı  $t$ 'de çevreden elde ettiği gözlemlere dayanarak gerçekleştirdiği eylemlerin sonucunda bir ödül elde etmektedir. Bu ödül fonksiyonu, yalnızca sistem hedeflerini değil, aynı zamanda ajanların bulunduğu iletişim modu ve kanal koşullarını da doğrudan hesaba katacak şekilde Eşitlik 4.26'da tanımlanmıştır:

$$r_j^t = -\kappa_1 F(p_j^t) - \left( \kappa_2 \frac{\zeta_j^t}{\zeta_j} \right) \delta(\theta_j - 1) - \left[ \kappa_3 A_j^t + \kappa_4 G(C_{j,I}^t - C_{j,I}^{min}) \right] \delta(\theta_j) - \left[ \kappa_5 A_j^t + \kappa_6 G(C_{j,S}^t - C_{j,S}^{min}) \right] \delta(\theta_j - 2) \quad (4.26)$$

Bu ödül fonksiyonu dört ana terimden oluşmaktadır. İlk terim, ajan tarafından o zaman adımında harcanan iletim gücü  $p_j^t$  'ye dayalı olarak belirlenen enerji tüketimini cezalandırmaktadır. Burada  $F(\cdot)$  genelleştirilmiş bir güç ceza fonksiyonu olup, tipik olarak doğrusal ya da logaritmik olarak tanımlanabilir. Bu terime bağlı katsayı olan  $\kappa_1$  , enerji tüketiminin toplam ödül üzerindeki ağırlığını belirler. İkinci terim, V2V haberleşme modu için tanımlanmıştır. Eğer ajan V2V modunda çalışıyorsa, elde edilen

veri miktarının hedeflenen minimum kapasite eşiği  $\zeta_j$  'ye göre oranı, ceza biçiminde modele dahil edilir. Bu yapı, V2V modunun kapasite bakımından yetersiz kaldığı durumları ödül düzeyinde olumsuz yansıtır. Üçüncü terim, V2I moduna özgüdür ve hem o andaki AoI değeri hem de iletişim kapasitesinin minimum eşik değerinin altına düşmesi durumlarını kapsayan çiftli bir ceza teriminden oluşur.  $\kappa_3$  ve  $\kappa_4$  sırasıyla bilgi tazeliği ve kapasite eksikliği durumlarına duyarlılığı kontrol ederken,  $G(\cdot)$  kapasite farkını işleyen pozitif tanımlı bir fonksiyondur. Benzer şekilde dördüncü terim, V2H haberleşme moduna geçilmesi durumunda aktive olur. Bu durumda da hem yüksek AoI değeri hem de yetersiz HAPS kapasitesi durumları ödül fonksiyonu içinde ceza olarak uygulanır. Bu bağlamda  $\kappa_5$  ve  $\kappa_6$ , V2H modundaki bilgi tazeliği ve kanal yeterliliğine verilen ağırlıkları yansıtır.

Eşitlik 4.26'daki  $\delta\{\cdot\}$  fonksiyonu, klasik Kronecker delta olup yalnızca içerideki koşul sağlandığında 1 değerini alır. Böylece her terim yalnızca ilgili mod aktifken hesaba katılır. Bu yapı, mod bağımlı ödül mekanizması oluşturur ve ajanların mod seçimlerini daha hassas şekilde optimize etmelerini sağlar.

Sonuç olarak Eşitlik 4.26, MADDPG algoritması kapsamında her ajan için çok boyutlu hedeflerin (enerji verimliliği, kapasite yeterliliği ve bilgi güncelliği) dengeleyici biçimde öğrenilebilmesine olanak tanır. Bu fonksiyon aracılığıyla ajanlar, yalnızca mevcut durumu iyileştirmekle kalmaz, aynı zamanda sistem genelindeki performansı optimize edecek politikalar geliştirir.

Eşitlik 4.27'de gösterilen fonksiyonun temel amacı, tüm müfrezelerin toplam kanal girişimini minimize ederek sistem genelinde daha kararlı ve düşük gecikmeli bir iletişim sağlamaktır. Logaritmik fonksiyonun kullanılması sayesinde, yüksek değerli girişimlerin ödül üzerindeki etkisi baskılanmakta ve böylece daha adil bir değerlendirme sağlanmaktadır. Girişimin ortalaması alınarak, tüm ajanlar arasında yük paylaşımı teşvik edilir ve iletişim ortamının dengeli bir şekilde kullanılması sağlanır.

$$r_g^t = -\frac{1}{P} \sum_{j \in P} \sum_{k \in K} \log_{10}(I_j^t[k]) \quad (4.27)$$

Eşitlik 4.28 ise MADDPG yaklaşımında ajanların politikalarını optimize etmek için kullanılan politika gradyan (policy gradient) güncellemesini açıklamaktadır.

$$\begin{aligned} \nabla_{\theta_j} J_j = & \mathbb{E}_{s,a \sim D} \left[ \nabla_{\theta_j} \pi_j(a_j | s_j) \nabla_{a_j} Q_{\Psi_1}^{g1}(s,a) \right] \\ & + \sum_{k=1}^M \mathbb{E}_{s_j, a_j \sim D} \left[ \nabla_{\theta_j} \pi_j(a_j | s_j) \nabla_{a_j} Q_{\Phi_{j,k}}^{j,k}(s_j, a_j) \right] \\ & + \mathbb{E}_{s_j, a_j \sim D} \left[ \nabla_{\theta_j} \pi_j(a_j | s_j) \nabla_{a_j} Q_{\Phi_H}^H(s_j, a_j) \right] \end{aligned} \quad (4.28)$$

Bu ifadeye üç farklı eleştiri (critic) fonksiyonundan yararlanılır.  $Q_{\Psi_1}^{g1}(s,a)$  küresel ödüle göre sistem performansını değerlendirirken,  $Q_{\Phi_{j,k}}^{j,k}(s_j, a_j)$  belirli bir kanal üzerindeki yerel ödül katkısını analiz eder,  $Q_{\Phi_H}^H(s_j, a_j)$  ise HAPS ile yapılan iletişimin ödül üzerindeki etkisini değerlendirir. Beklenen değer operatörleri  $\mathbb{E}$ , deneyim tekrar havuzu (experience replay buffer) üzerinden örneklenen durum-eylem çiftleri üzerinden alınır. Böylece ajanlar, kendi politikalarını hem yerel hem küresel düzeyde optimize ederek güç tüketimini azaltırken AoI artırmayı amaçlarlar.

Bu çalışmada önerilen HAPS destekli V2X sisteminde kaynak yönetiminin çevrim içi ve otonom biçimde optimize edilmesi için MADDPG algoritması kullanılmaktadır. Aşağıda sunulan Algoritma 1, çok ajanlı ortamda her bir ajanın politika fonksiyonunu, hem küresel hem de yerel eleştirmen (critic) yapıları ile nasıl güncellediğini adım adım göstermektedir. Öğrenme süreci, klasik aktör-eleştirmen mimarisi üzerine kurulmuş olup, her ajan kendi yerel gözlemlerine dayalı olarak aksiyon üretirken; merkezi bir yapı üzerinden ödüller ve geri bildirimler değerlendirilir. Küresel ödül, sistem çapındaki girişim miktarını en aza indirmeye çalışırken; yerel ödüller, her ajanın kendi AoI, kapasite yeterliliği ve güç tüketimine dayalı performansını optimize etmesini sağlar. Algoritma, hedef ağ güncellemeleri ve deneyim tekrarları (experience replay) yoluyla kararlılığı artırmakta ve öğrenilen politikaların zaman içinde yakınsaklık göstermesini sağlamaktadır. Bu yapının simülasyon ortamına entegre edilmesiyle, ajanlar veri güncelliğini yüksek tutarken aynı zamanda enerji ve spektrum kaynaklarını da daha verimli kullanmayı öğrenmektedir. Burada  $y_{\Psi_j}^g = r_j^g + \gamma \min Q_{\Psi_j}^g(s', a')$  küresel eleştirmen için hedef Q-değerinin hesaplandığı yerdir.

$L(\psi_i)$  küresel eleştirmen ağı  $Q_\psi^g$  için kullanılan ortalama kare hata tabanlı kayıp fonksiyonudur.  $\nabla J_{\theta_j}$  aktör ağının (politika fonksiyonunun) güncellenmesini sağlayan politika gradyan ifadesidir. Son ifade ise hedef ağların yumuşak güncellenmesini ifade eder. Bu yöntem, ani parametre değişikliklerinin öğrenme sürecinde oluşturabileceği dengesizlikleri engeller ve politikaların zaman içinde daha kararlı biçimde yakınsamalarını sağlar.

### **Algoritma 1: MADDPG Optimizasyonu**

Ana küresel eleştirmen ağlarını  $Q_\psi^u$  ve  $Q_\psi^v$  olarak başlat.

Hedef küresel eleştirmen ağlarını  $Q_\psi^u$  ve  $Q_\psi^v$  olarak başlat.

Her bölüm (episode) için HAPS ve araç konumlarını güncelle, ilgili kazancı hesapla.

Yüklü sıralar,  $\zeta^j$ 'yi başlat ve maksimum teslim süresini  $T = 100$  ms olarak ayarla.

Her zaman adımı  $t$  için her ajan  $i$  için, zamanın gözlemleri  $s_t^i$  ve eylemi  $a_t^i = \pi_i(s_t^i)$  seç. Küresel durum-aksiyon çifti  $s_t = [s_t^1, \dots, s_t^n]$  ve eylem vektörü  $a_t = [a_t^1, \dots, a_t^n]$  olarak belirle.

Küresel ve yerel ödülleri al,  $r_t^i$  ve  $r_t^j$ . Durum geçişlerini tanımla ve tekrar oynatma belleğine  $D$  ekle. Küresel eleştirmen güncellemesi: Bellekten  $S$  boyutunda bir mini-toplu örnek al.

$$\hat{y}_{\psi_i} = r_j^u + \gamma \min_{a'} Q_{\psi_i}^u(s', a') \text{ olarak ayarla.}$$

Aşağıdaki kayıp fonksiyonunu minimize ederek küresel eleştirmeni güncelle:

$$L(\psi_i) = (1/S) \sum_j (Q_{\psi_i}^u(s_j', a_j') - \hat{y}_{\psi_i})^2$$

Hedef parametrelerini güncelle:  $\psi_i' \leftarrow \tau \psi_i + (1 - \tau) \psi_i'$

Her bölüm numarası  $d$ 'ye tam bölünüyorsa her ajan  $i$  için  $\hat{y}_{\phi_i} = r_j^v + \gamma Q_{\phi_i}^v(s', a')$  olarak ayarla.

Yerel eleştirmeni şu kayıp fonksiyonunu minimize ederek güncelle:

$$L(\phi_i) = (1/S) \sum_j (Q_{\phi_i}^v(s_j', a_j') - \hat{y}_{\phi_i})^2$$

Yerel aktörü maksimize ederek güncelle:

$$\nabla_{\theta_i} = (1/S) \sum_j [\nabla_{a_i} \pi_i(a_i \vee s_i) \nabla_{a_i} Q_{\psi_i}^u(s^a, a) + \nabla_{a_i} \pi_i(a_i \vee s_i) \nabla_{a_i} Q_{\phi_i}^v(s_i', a_i)]$$

Hedef ağ parametrelerini güncelle:

$$\theta_i' \leftarrow \tau \theta_i + (1 - \tau) \theta_i'$$

$$\phi_i' \leftarrow \tau \phi_i + (1 - \tau) \phi_i'$$

## 5. SONUÇLAR VE ÖNERİLER

Bu bölümde, önerilen MADDPG tabanlı çerçevenin HAPS destekli bir V2X haberleşme ortamındaki performansı değerlendirilmektedir. Simülasyon senaryosu, AoI, haberleşme verimliliği ve enerji tüketimi gibi temel performans metriklerini optimize etmek amacıyla dinamik koşullar altında çalışan çok etmenli bir sistemdeki araç platonlarını (PL'leri) dikkate almaktadır.

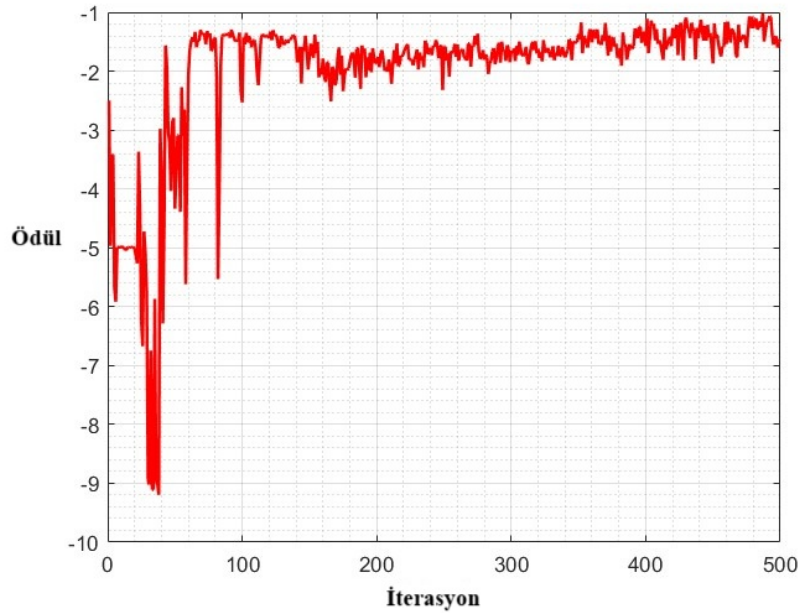
### 5.1. Sonuçlar

İletişim mesafesi, iletim gücü ve ağ konfigürasyonları gibi simülasyon parametreleri Çizelge 5.1'de özetlenmiştir. Bu parametreler, tüm deneyler için temel yapılandırmayı oluşturarak sistemin farklı koşullar altındaki uyarlanabilirliğini ve performansını analiz etmek için kapsamlı bir çerçeve sunmaktadır.

Çizelge 5.1. Başlangıç Parametreleri

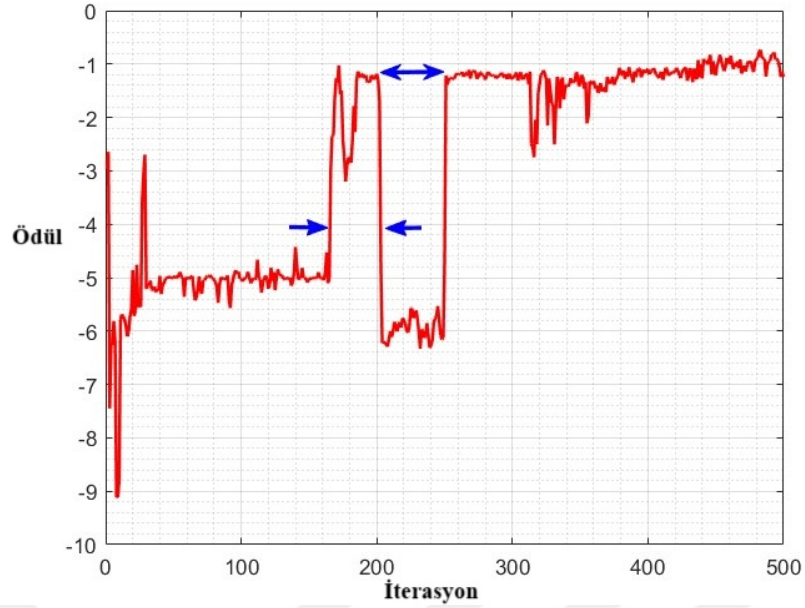
Parametre Açıklaması	Değer
Takım araçları arasındaki iletişim mesafesi	25 metre
Takım liderinin maksimum iletim gücü	30 dBm
V2I iletişimi için gereken minimum veri hızı	540 kbps
Mevcut toplam bant genişliği	180 kHz
Araçlar arası iletişimde kullanılan veri boyutu	4000 Bayt
Eğitimde kullanılan yığın (batch) boyutu	64
Tekrar belleğinin boyutu	50000
Ödül indirim faktörü ( $\gamma$ )	0.99
Gürültü seviyesinin standart sapması ( $\sigma$ )	0.3 dB
Aktör katman boyutları (Birinci katman, İkinci katman)	[1024, 512]
Eleştirmen (critic) katman boyutları	[1024, 512, 256]
Hedef ağlar için ağırlık güncelleme oranı ( $\tau$ )	0.005
Mevcut kaynak bloklarının sayısı	3
Her bir filodaki araç sayısı	6
Toplam filo sayısı	5
HAPS yüksekliği	20 km
Kapasite üst sınırı	1 Mbps
Aktör öğrenme hızı ( $\alpha$ )	0.0001
Eleştirmenin öğrenme oranı ( $\beta$ )	0.001

İlk olarak RL süreci boyunca PL'ler için ödül fonksiyonunun evrimi Şekil 5.1-5.5 arasında gösterilmiştir. Şekil 5.1, PL-1'in başlangıçta çevresini keşfettiğini ve bu nedenle ödül fonksiyonunun düşük ve dalgalı seyrettiğini göstermektedir. Ayrıca, ajan çevresini daha iyi tanımaya başladıkça ödül değerinin arttığı ve dalgalanmaların azaldığı gözlemlenmiştir; bu da ajanın daha tutarlı bir şekilde performans göstermeye başladığını ifade etmektedir. 200. döngüden itibaren ödül fonksiyonunun -1 değerine yakınsadığı açıkça görülmekte olup, bu yakınsama 400. döngüden sonra daha da belirgin hâle gelmektedir. Bu durum, ilgili ajanın optimal karar verme politikasına ulaştığını ve performansını maksimize ettiğini göstermektedir.



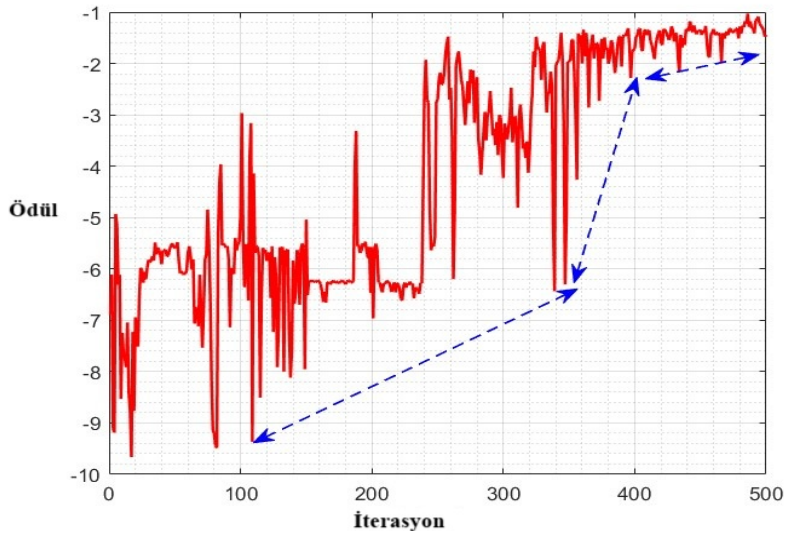
Şekil 5.1. Ajan-1 Ödül Fonksiyonu

Şekil 5.2, PL-2 için öğrenme sürecinin farklı evrelerini ortaya koymaktadır. İlk 160 döngü boyunca ödül değeri düşük seviyelerde dalgalı bir seyir izlemekte olup, bu durum ajanın başlangıçta keşif ve öğrenme aşamasında olduğunu göstermektedir. 160 ila 200. döngüler arasında ödül değerinin -1 seviyesine yaklaşmaya çalıştığı görülmekle birlikte, 200. döngüde ani bir düşüş yaşanarak ödül değeri tekrar düşük seviyelere gerilemiştir. Neyse ki bu düşüş uzun sürmemiş ve 260. döngüden itibaren daha kararlı bir şekilde -1 seviyesine yakınsama sağlanmıştır. Bu tür ani değişimler, çevresel koşullarda ya da ajanın öğrenme stratejisinde meydana gelen belirli ayarlamaların bir sonucu olabilir.



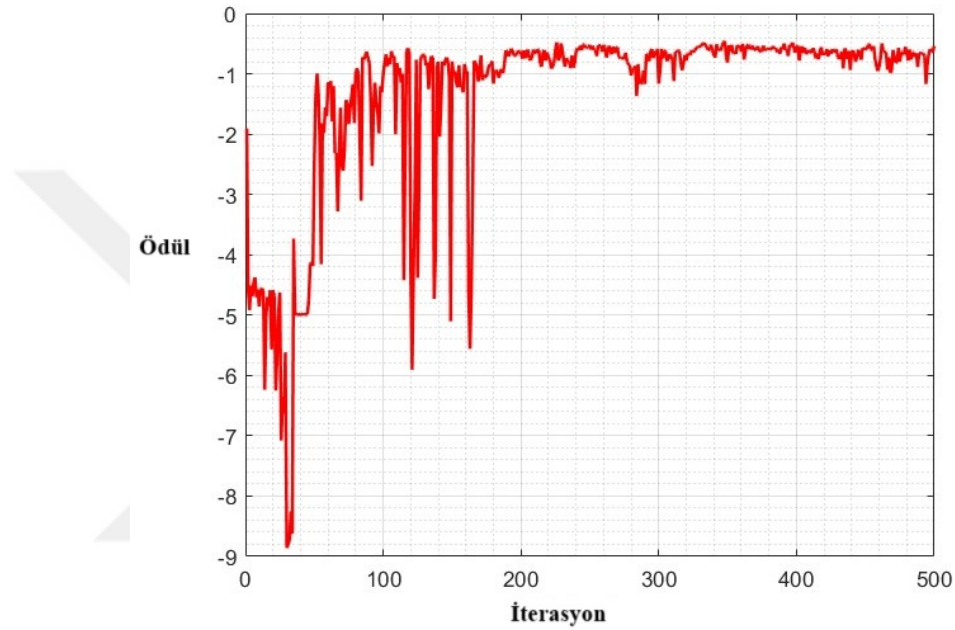
Şekil 5.2. Ajan-2 Ödül Fonksiyonu

Şekil 5.3'te, önceki şekillerde gözlemlenenden daha uzun ve daha belirgin dalgalanmalar görülmektedir. İlk 360 döngü boyunca yaşanan kararsız ödül fonksiyonu, PL-3'ün çevreyi keşfetme ve optimal eylemleri öğrenme sürecinde daha fazla zorluk yaşadığını göstermektedir. Ancak 350. döngü itibaren ödül fonksiyonunda belirgin bir iyileşme gözlemlenmiş ve ajan kısa sürede daha yüksek ödül seviyelerine ulaşmıştır. Bu artış, ajanın öğrenme sürecinde kritik bir dönüm noktasına ulaştığını ve artık optimal politikalara daha yakın kararlar aldığını göstermektedir. 400. döngüden sonra ödül değeri -1 seviyesinde istikrar kazanmış ve ajanın performansı yüksek düzeyde bir kararlılığa ulaşmıştır.



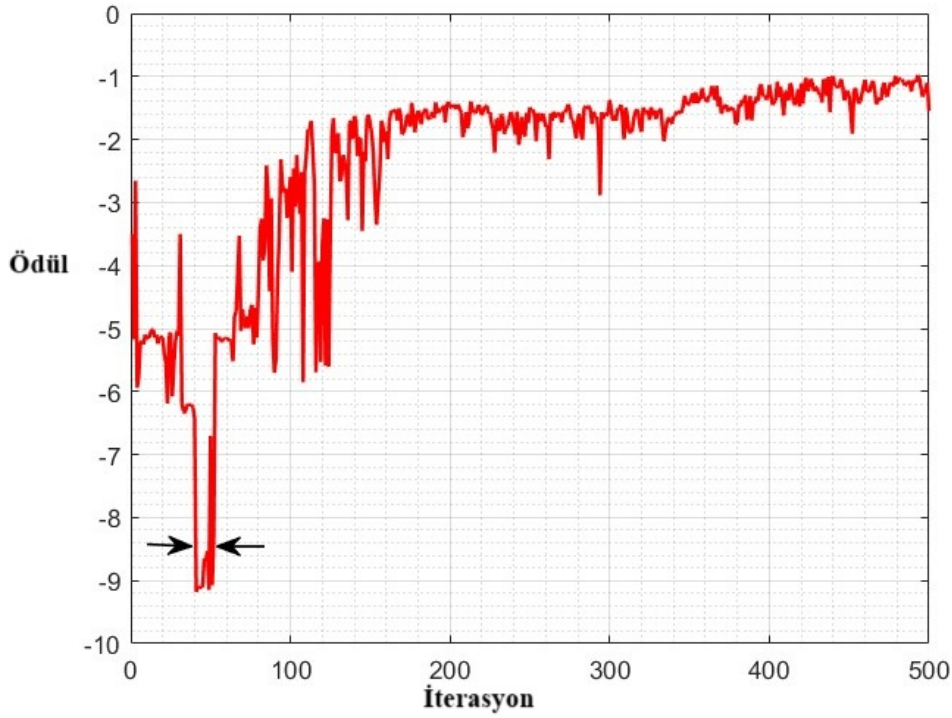
Şekil 5.3. Ajan-3 Ödül Fonksiyonu

Şekil 5.4'ten görüldüğü üzere, PL-4'ün çevreyi keşfetme ve öğrenme süreci 160 döngü sürmektedir. İlk 50 döngü boyunca ödül değeri oldukça düşük seviyelerde dalgalanmış ve yaklaşık olarak -5 ile -8 arasında seyretmiştir. 50. döngüden sonra ödül fonksiyonu hızlı bir şekilde yükselmiş ve döngü sayısının artmasıyla birlikte -1 seviyesinde istikrar kazanmaya çalışmıştır. 160. döngüden sonra ise ödül fonksiyonu daha kararlı bir görünüm sergilemeye başlamış, dalgalanmalarda önemli ölçüde azalma gözlenmiştir.



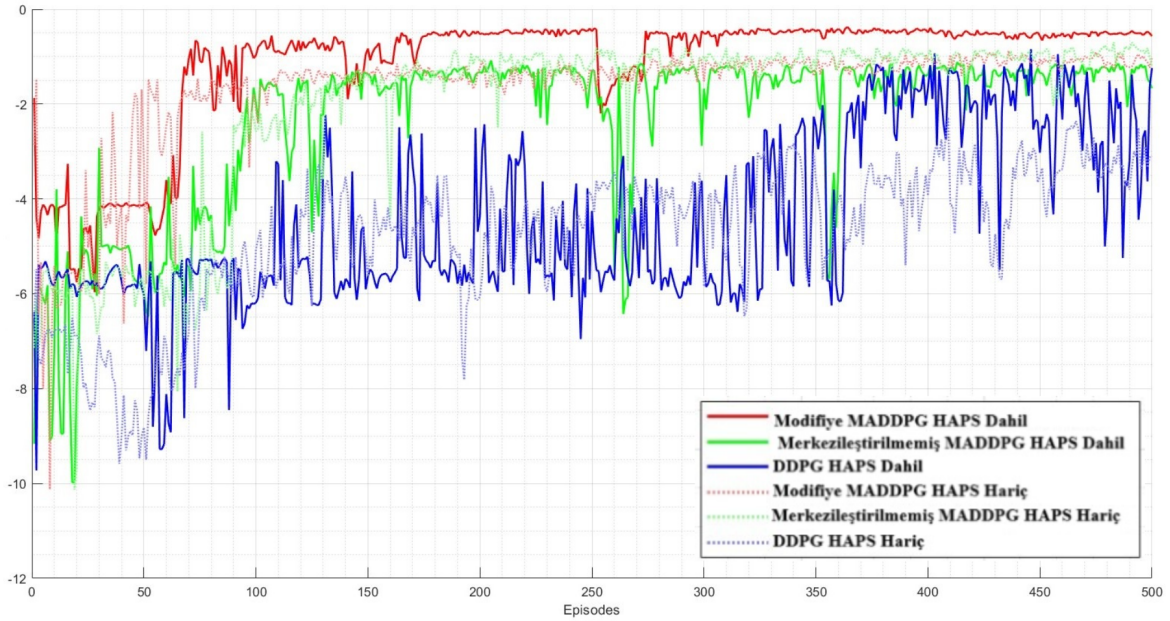
Şekil 5.4. Ajan-4 Ödül Fonksiyonu

Şekil 5.5, öğrenme süreci boyunca PL-5 ajanının performansını yansıtan ödül fonksiyonundaki değişimi göstermektedir. İlk 60 döngü boyunca ödül değerinin oldukça düşük olduğu ve yaklaşık olarak -9 ile -5 arasında dalgalandığı görülmektedir. Bu durum, ajanın çevresini tanımaya çalıştığı ve henüz optimal kararlar veremediği başlangıç aşamasını yansıtmaktadır. 60. döngüden itibaren ise ödül değeri kademeli olarak artmaya başlamış ve 160. döngü civarında yaklaşık -2 seviyesine ulaşmıştır. Bu durum, ajanın çevreye uyum sağladığını ve daha etkili kararlar vermeye başladığını göstermektedir. 160. döngüden sonra ise ödül fonksiyonundaki dalgalanmalar azalmış ve 200. döngüden itibaren yaklaşık -1 değeri etrafında kararlı bir yapıya ulaşılmıştır. Bu son yakınsama, ajanın pekiştirmeli öğrenme sürecini başarıyla tamamladığını göstermektedir.



Şekil 5.5. Ajan-5 Ödül Fonksiyonu

Şekil 5.6, bu çalışmada önerilen modifiye MADDPG algoritmasının ödül fonksiyonunun yakınsama performansını; tüm karar verme süreçlerinin RSU'da gerçekleştirildiği tam merkezî geleneksel DDPG algoritması ve RSU'nun sürece dahil edilmediği tamamen dağıtık MADDPG algoritması ile karşılaştırmaktadır. Ayrıca, bu karşılaştırmalara HAPS desteğinin V2X sistemlerine etkisini göstermek amacıyla, algoritmaların HAPS destekli ve destekli versiyonları da dâhil edilmiştir. Bu şekilden çıkarılabilecek ilk sonuç, HAPS desteğinin tüm algoritmaların öğrenme performansını genel olarak artırdığıdır. Bununla birlikte, HAPS destekli modifiye MADDPG algoritmasının, en yüksek ödül seviyesine ulaşan yapı olduğu açıkça görülmektedir. Bu algoritma HAPS desteği olmadan çalıştırıldığında ise öğrenme süreci çok düşük ödül değerleriyle başlamaktadır; hatta ödül değeri -10'a kadar düşebilmektedir. 100. döngü sonrasında ödül değeri yaklaşık -1 seviyesinde yakınsamaya ulaşsa da, HAPS destekli versiyona kıyasla hâlâ daha düşük performans sergilemektedir.

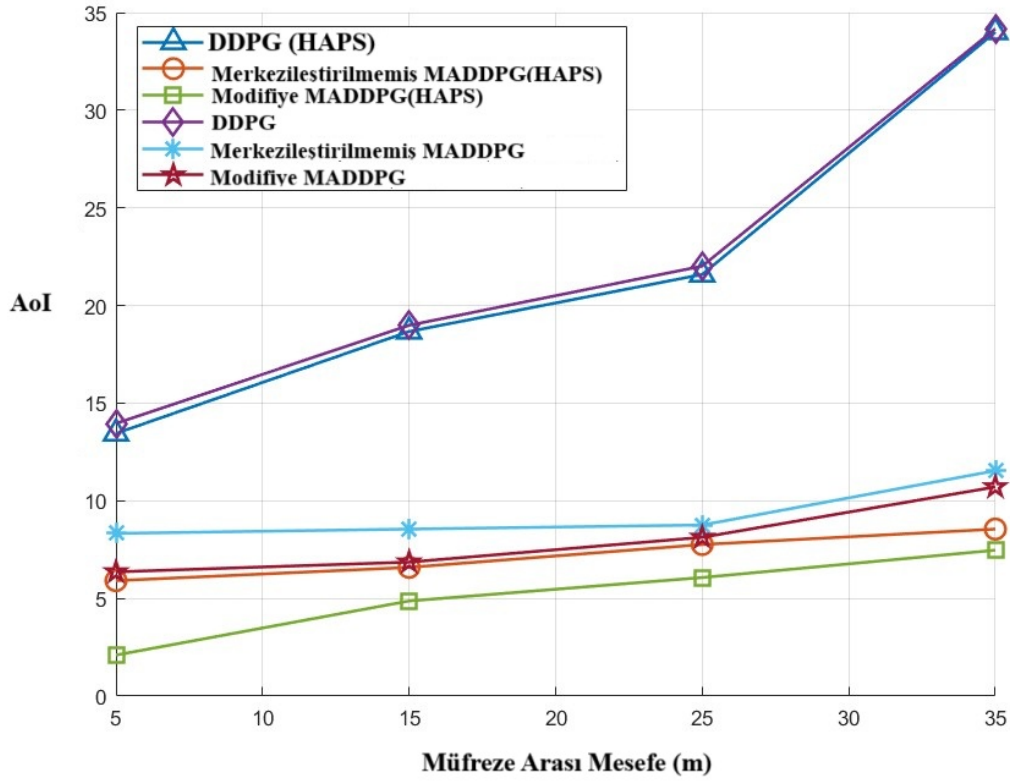


Şekil 5.6. Algoritmaya göre ödül fonksiyonları yakınsamalarının karşılaştırılması.

Modifiye MADDPG, klasik MADDPG'nin geliştirilmiş bir versiyonu olup daha kararlı ve hızlı yakınsama amaçlanırken; merkezileştirilmemiş yapı, öğrenme sürecinin her ajan tarafından yerel olarak yürütüldüğü dağıtık bir yaklaşımdır. DDPG ise tek ajanlı ortamlarda yaygın olan, klasik bir derin pekiştirmeli öğrenme yöntemidir. Şekil 5.6'da yeşil renkle gösterilen tamamen dağıtık MADDPG algoritması kullanıldığında ise, ödül fonksiyonundaki ani yükseliş ve düşüşlerin 340. döngüye kadar devam ettiği görülmektedir. Bu durum, öğrenme sürecinin modifiye MADDPG algoritmasına kıyasla daha yavaş ilerlediğini göstermektedir. Ayrıca, HAPS entegrasyonu bu yapının performansına anlamlı bir katkı sağlamamıştır. Öte yandan, mavi renkle gösterilen klasik DDPG algoritması, incelenen tüm yapılar arasında en düşük performansı sergilemektedir. HAPS desteğiyle birlikte bu algoritmanın ödül değeri zamanla yükseliyor olsa da, ödül fonksiyonunun yakınsama sürecindeki ciddi dalgalanmalar 500 döngü sonunda dahi ortadan kaldırılamamaktadır. Bu durum, geleneksel DDPG algoritmasının çok etmenli sistemler için gereken esnekliği ve adaptasyonu sağlayamadığını açıkça göstermektedir.

Buna ek olarak, Şekil 5.7'de algoritmaların AoI performanslarını, araçlar arası mesafeye bağlı olarak karşılaştırmaktadır. Burada daha düşük AoI değerlerinin, daha güncel bilgiye işaret ettiği unutulmamalıdır. Şekilden görüldüğü üzere, araçlar arası

mesafe arttıkça tüm algoritmalar için AoI değeri de artmaktadır, yani bilgi güncelliği azalmaktadır.



Şekil 5.7. Algoritmaya göre AoI metriği karşılaştırılması

Şekil 5.7'ye göre özellikle klasik DDPG algoritması, HAPS destekli olsun ya da olmasın, her mesafede en yüksek AoI değerlerini üretmektedir. Örneğin 35 metrelik bir mesafe için, tamamen dağıtık ve modifiye edilmiş MADDPG algoritmaları 13 ms'nin altında AoI değerlerine ulaşırken, klasik DDPG algoritmasında bu değer yaklaşık 35 ms civarındadır. Ayrıca şekilden çıkarılabilecek bir diğer önemli sonuç, araçlar arası mesafe arttıkça klasik DDPG algoritmasının ürettiği AoI değerinin daha keskin biçimde yükselmesidir. Örneğin, mesafe 5 metreden 35 metreye çıkarıldığında klasik DDPG'nin ortalama AoI değeri yaklaşık 20 ms artarken; modifiye ve dağıtık MADDPG algoritmalarında bu artış 6 ms'yi geçmemektedir. Bu bulgular, klasik DDPG algoritmasının HAPS desteği bulursa dahi bilgi güncelleme hızının görece yavaş olduğunu ve bu nedenle dinamik V2X senaryoları için uygun olmadığını göstermektedir.

Tamamen dağıtık MADDPG algoritması ise klasik DDPG'ye kıyasla daha iyi performans göstermektedir. Bu algoritmanın HAPS destekli versiyonu, tüm mesafeler için 10 ms'nin altında AoI değeri üretmektedir. Desteksiz versiyon ise, yaklaşık 3 ms daha yüksek AoI değeri göstermektedir. Bu sonuçlar, HAPS desteğinin dağıtık MADDPG algoritmasının bilgi güncelleme performansını artırdığını ve AoI değerini düşürdüğünü açıkça ortaya koymaktadır.

Modifiye MADDPG algoritmasının HAPS'siz versiyonunun, performans açısından HAPS destekli tamamen dağıtık MADDPG algoritmasına oldukça yakın sonuçlar verdiği gözlenmiştir. Ancak en düşük AoI değerleri (yani en güncel bilgiye sahip yapı) HAPS destekli modifiye MADDPG algoritmasında elde edilmiştir. Bu nedenle, HAPS ile entegre edilmiş modifiye MADDPG algoritmasının hem bilgi tazeliği açısından hem de öğrenme verimliliği bakımından en başarılı sonuçları sunduğu ve önerilen V2X senaryosu için en uygun yöntem olduğu açıkça görülmektedir.

## 5.2. Öneriler

Bu tez çalışmasında, HAPS destekli V2X haberleşme sistemleri için kaynak yönetimini optimize etmeye yönelik çok etmenli bir yaklaşım geliştirilmiş ve önerilen çözüm çerçevesi, MADDPG algoritması ile gerçekleştirilmiştir. Sistemin hedefleri; bilgi güncelliğini ifade eden AoI değerlerini minimize etmek, spektrum kaynaklarını etkin bir şekilde kullanmak ve enerji tüketimini azaltmak olarak belirlenmiştir. Bu doğrultuda geliştirilen matematiksel model, hem sürekli hem de ayrık eylem uzaylarını kapsayan karmaşık bir optimizasyon problemi olarak tanımlanmış ve çözüm için pekiştirmeli öğrenme tabanlı bir yöntem benimsenmiştir.

Yürütülen simülasyon çalışmaları, önerilen HAPS destekli MADDPG algoritmasının, klasik DDPG ve tamamen dağıtık MADDPG algoritmalarına kıyasla hem öğrenme sürecinde daha hızlı yakınsadığını hem de daha kararlı bir politika öğrenimi gerçekleştirdiğini göstermektedir. Özellikle HAPS entegrasyonu sayesinde, bilgi güncelliği önemli ölçüde artırılmış ve AoI değerlerinde ciddi iyileşmeler sağlanmıştır. Aynı zamanda sistemin enerji verimliliği de korunmuş ve iletişim performansında anlamlı artışlar elde edilmiştir. Elde edilen sonuçlar, HAPS'in yüksek

kapsama alanı ve doğrudan görüş hattı avantajlarıyla, karasal haberleşme altyapılarını tamamlayıcı önemli bir bileşen olarak kullanılabileceğini ortaya koymaktadır.

Bununla birlikte, önerilen yöntem mevcut yapının bazı sınırlılıklarını da içermektedir. Örneğin, öğrenme süreci yalnızca simülasyon ortamında gerçekleştirilmiş olup, gerçek zamanlı uygulamalara geçildiğinde çevresel değişkenlikler ve fiziksel donanım kısıtları daha karmaşık etkilere yol açabilir. Ayrıca, ajanların eğitim süreci, RSU ile desteklenmiştir; bu yapı, her zaman tam dağıtık bir sistemin avantajlarını tam olarak yansıtmayabilir.

Önerilen model, gerçek trafik verileri ve dinamik kanal koşullarıyla test edilerek daha gerçekçi ortamlara entegre edilebilir. MADDPG algoritması yerine, daha yeni ve veri verimli alternatif pekiştirmeli öğrenme algoritmaları (örneğin, Tamamen merkezileştirilmemiş MADDPG (Fuly Decentralized MADDPG, FD-MADDPG)) ile karşılaştırmalı çalışmalar yapılabilir. Enerji verimliliğini artırmaya yönelik yeni ödül fonksiyonları geliştirilebilir. Uygulama tarafında, HAPS platformlarının hareketliliği ve görev süresi gibi parametreler modele dâhil edilerek daha esnek ve ölçeklenebilir sistem tasarımları gerçekleştirilebilir.

## KAYNAKLAR

- Abbasi, O., Yadav, A., Yanıkömeroğlu, H., Đào, N. D., Senarath, G. ve Zhu, P., 2024, HAPS for 6G networks: Potential use cases, open challenges, and possible solutions, *IEEE Wireless Communications*, 31 (3), 324–331.
- Alfattani, S., Jaafar, W., Hmamouche, Y., Yanikomeroğlu, H. ve Yongacoglu, A., “Link budget analysis for reconfigurable smart surfaces in aerial platforms”, *IEEE Open Journal of the Communications Society*, 2 (2021): 1980–1995.
- Alwarafy, A., Abdallah, M., Ciftler, B. S., Al-Fuqaha, A. ve Hamdi, M., 2022, The frontiers of deep reinforcement learning for resource management in future wireless HetNets: Techniques, challenges, and research directions, *IEEE Open Journal of the Communications Society*, 3, 322–365.
- Arena, F., Pau, G. ve Severino, A., 2020. A review on IEEE 802.11p for intelligent transportation systems. *Journal of Sensor and Actuator Networks*, 9(2).
- Arulkumar, K., Deisenroth, M. P., Brundage, M. ve Bharath, A. A., 2017, Deep reinforcement learning: A brief survey, *IEEE Signal Processing Magazine*, 34 (6), 26–38.
- Avcil, M. N., Soyuturk, M. ve Kantarci, B., 2024, Fair and efficient resource allocation via vehicle-edge cooperation in 5G-V2X networks, *Vehicular Communications*, 48, 100773.
- Bazzi, A., Berthet, A. O., Campolo, C., Masini, B. M., Molinaro, A. ve Zanella, A., 2021, On the design of sidelink for cellular V2X: A literature review and outlook for future, *IEEE Access*, 9, 97953–97980.
- Choi, C. S., 2022. Modeling and analysis of priority-based distributed access control in vehicular networks. *IEEE Transactions on Vehicular Technology*, 71(6), ss. 6848–6852.
- Clancy, J., Mullins, D., Deegan, B., Horgan, J., Ward, E., Eising, C., Denny, P., Jones, E. ve Glavin, M., “Wireless access for V2X communications: Research, challenges and opportunities”, *IEEE Communications Surveys & Tutorials*, Vol. 26, No. 1, pp. 123-156, (2024).
- Ernst, R., Stöhrmann, D., Bendrick, A. ve Kostrzewa, A., 2023, Application-centric network management – addressing safety and real-time in V2X applications, *ACM Transactions on Embedded Computing Systems*, 22 (2), 1–25.
- Gomez-Barquero, D., Gimenez, J. J. ve Beutler, R., 2020. 3GPP enhancements for television services: LTE-based 5G terrestrial broadcast. *Wiley Encyclopedia of Electrical and Electronics Engineering*.
- Giordani, M. ve Zorzi, M., 2020, Non-terrestrial networks in the 6G era Challenges and opportunities, *IEEE Network*, 35 (2), 244–251.

- Gyawali, S., Xu, S., Qian, Y. ve Hu, R. Q., 2020, Challenges and solutions for cellular based V2X communications, *IEEE Communications Surveys & Tutorials*, 23 (1), 222–255.
- Hasan, M. K., Jahan, N., Nazri, M. Z. A., Islam, S., Khan, M. A., Alzahrani, A. I., Alalwan, N. ve Nam, Y., 2024, Federated learning for computational offloading and resource management of vehicular edge computing in 6G-V2X network, *IEEE Transactions on Consumer Electronics*, 70 (1), 3827–3847.
- Huang, Y., 2020, Deep Q-networks, *Deep Reinforcement Learning: Fundamentals, Research and Applications*, 135–160.
- Jang, B., Kim, M., Harerimana, G. ve Kim, J. W., 2019, Q-learning algorithms: A comprehensive classification and applications, *IEEE Access*, 7, 133653–133667.
- İnce, A. M., Canbilen, A. E. ve Yanıkömeroğlu, H., 2024, HAPS-Enabled V2X Architecture for Hyper Reliable and Low-Latency Communication (HRLLC) in 6G Networks, *2024 6th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, IEEE, 1–6, Temmuz 2024.tworks, *2025 17th International Conference on Communication Systems and Networks (COMSNETS)*, IEEE, 324–331.
- Kenney, J. B., 2011, Dedicated short-range communications (DSRC) standards in the United States, *Proceedings of the IEEE*, 99 (7), 1162–1182.
- Kurt, G. K., Khoshkholgh, M. G., Alfattani, S., Ibrahim, A., Darwish, T. S., Alam, M. S. ve Yongaoğlu, A., 2021, A vision and framework for the high altitude platform station (HAPS) networks of the future, *IEEE Communications Surveys & Tutorials*, 23 (2), 729–779.
- Lai, X., Yang, Z., Xie, J. ve Liu, Y., 2024, Reinforcement learning in transportation research: frontiers and future directions, *Multimodal Transportation*, 2024.
- Li, X., Lu, L., Ni, W., Jamalipour, A., Zhang, D. ve Du, H., 2022, Federated multi-agent deep reinforcement learning for resource allocation of vehicle-to-vehicle communications, *IEEE Transactions on Vehicular Technology*, 71 (8), 8810–8824.
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. ve Wierstra, D., 2015, Continuous control with deep reinforcement learning, *arXiv preprint*, arXiv:1509.02971.
- Liu, Z., Han, Y., Fan, J., Zhang, L. ve Lin, Y., 2020, Joint optimization of spectrum and energy efficiency considering the C-V2X security: A deep reinforcement learning approach, *2020 IEEE 18th International Conference on Industrial Informatics (INDIN)*, Cilt 1, IEEE, 315–320.
- Lou, Z., Belmekki, B. E. Y. ve Alouini, M. S., 2023, HAPS in the non-terrestrial network nexus: Prospective architectures and performance insights, *IEEE Wireless Communications*, 30 (6), 52–58.

- Luong, N. C., Hoang, D. T., Gong, S., Niyato, D., Wang, P., Liang, Y.-C. ve Kim, D. I., 2019, Applications of deep reinforcement learning in communications and networking: A survey, *IEEE Communications Surveys & Tutorials*, 21 (4), 3133–3174.
- Maglogiannis, V., Naudts, D., Hadiwardoyo, S., Van Den Akker, D., Marquez-Barja, J. ve Moerman, I., 2021, Experimental V2X evaluation for C-V2X and ITS-G5 technologies in a real-life highway environment, *IEEE Transactions on Network and Service Management*, 19 (2), 1521–1538.
- Nair, A. ve Tanwar, S., 2024, Resource allocation in V2X communication: State-of-the-art and research challenges, *Physical Communication*, 2024, 102351.
- Parvini, M., Javan, M. R., Mokari, N., Abbasi Arand, B. ve Jorswieck, E. A., 2021, AoI aware radio resource management of autonomous platoons via multi agent reinforcement learning, *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, IEEE, 1–6.
- Parvini, M., Javan, M. R., Mokari, N., Abbasi, B. ve Jorswieck, E. A., 2023, AoI-aware resource allocation for platoon-based C-V2X networks via multi-agent multi-task reinforcement learning, *IEEE Transactions on Vehicular Technology*, 72 (8), 9880–9896.
- Peng, H. ve Shen, X., 2020, Multi-agent reinforcement learning based resource management in MEC- and UAV-assisted vehicular networks, *IEEE Journal on Selected Areas in Communications*, 39 (1), 131–141.
- Saad, M. M., Tariq, M. A., Ajmal, M., Kim, D. ve Srivastava, G., 2025. Federated multi-agent reinforcement learning for resource allocation in NR-V2X mode 2. *IEEE Internet of Things Journal*.
- Saad, M. M., Tariq, M. A., Khan, M. T. R. ve Kim, D., 2024. Non-terrestrial networks: An overview of 3GPP Release 17 & 18. *IEEE Internet of Things Magazine*, 7(1), ss. 20-26.
- Saxena, C., Mukherjee, D., Thakur, K. P., Behera, S. ve Palit, B., “A Lightweight QoS-Aware Resource Allocation Method for NR-V2X Networks”, 2025 17th International Conference on Communication Systems and Networks (COMSNETS), pp.324–331, Jan. 2025.
- Sehla, K., Nguyen, T. M. T., Pujolle, G. ve Velloso, P. B., 2022, Resource allocation modes in C-V2X: from LTE-V2X to 5G-V2X, *IEEE Internet of Things Journal*, 9 (11), 8291–8314.
- Seid, A. M. ve Erbad, A., 2023, Multi-agent RL for SDN-based resource allocation in HAPS-assisted IoV networks, *ICC 2023 – IEEE International Conference on Communications*, IEEE, 1664–1669.

- Sze, V., Chen, Y.-H., Yang, T.-J. ve Emer, J. S., 2017, Efficient processing of deep neural networks: A tutorial and survey, *Proceedings of the IEEE*, 105 (12), 2295–2329.
- Trabelsi, A. D., Marouane, H., Zarai, F. ve Meddeb-Makhlouf, A., 2019, Dynamic scheduling algorithm based on priority assignment for LTE-V2X vehicular networks, *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, IEEE, 483–488.
- Vu, H. V., Farzanullah, M., Liu, Z., Nguyen, D. H. N., Morawski, R. ve Le-Ngoc, T., 2022, Multi-agent reinforcement learning for channel assignment and power allocation in platoon-based C-V2X systems, *2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring)*, IEEE, 1–5.
- Wang, J., Bai, L., Fang, Z., Han, R., Wang, J. ve Choi, J., 2024, Age of information based URLLC transmission for UAVs on pylon turn, *IEEE Transactions on Vehicular Technology*, 73 (6), 8797–8809.
- Ye, H., Li, G. Y. ve Juang, B.-H. F., 2019, Deep reinforcement learning based resource allocation for V2V communications, *IEEE Transactions on Vehicular Technology*, 68 (4), 3163–3173.
- Yoshizawa, T., Singelée, D., Muehlberg, J. T., Delbruel, S., Taherkordi, A., Hughes, D. ve Preneel, B., 2023, A survey of security and privacy issues in V2X communication systems, *ACM Computing Surveys*, 55 (9), 1–36.
- Yuan, Y., Zheng, G., Wong, K.-K. ve Letaief, K. B., 2021, Meta-reinforcement learning based resource allocation for dynamic V2X communications, *IEEE Transactions on Vehicular Technology*, 70 (9), 8964–8977.
- Zhang, Y., Mou, Z., Gao, F., Jiang, J., Ding, R. ve Han, Z., 2020, UAV-enabled secure communications by multi-agent deep reinforcement learning, *IEEE Transactions on Vehicular Technology*, 69 (10), 11599–11611.
- Zhang, X., Peng, M., Yan, S. ve Sun, Y., 2019, Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications, *IEEE Internet of Things Journal*, 7 (7), 6380–6391.
- Zhang, X., Zhang, H., Tang, H., Liang, L., Cheng, L., Chen, X., Ding, W. ve Zhang, X.-P., “A scalable mean-field MARL framework for multi-objective V2X resource allocation”, *IEEE Transactions on Intelligent Vehicles*, 2024. DOI: 10.1109/TIV.2024.3422506.