

Yakıt Hücreli Hibrit Elektrikli Aracın Derin Takviyeli Öğrenme ile Enerji Yönetimi Stratejisi

Energy Management Strategy of Fuel Cell Hybrid Electric Vehicle via Deep Reinforcement Learning

Tuna Bilgin¹, Batu Gürses², Saykut İzgu³, Ali Kerem Yalçın⁴, Sabri Yılmaz⁵

^{1,2,3,4,5} AVL Araştırma ve Mühendislik, İstanbul, Türkiye

tuna.bilgin@avl.com
batu.gurses@avl.com
saykut.izgu@avl.com
alikerem.yalcin@avl.com
sabri.yilmaz@avl.com

Özetçe

Bu çalışmada yakıt hücresi yükünü ve bataryaya sahip bir araç modeli kullanılarak sistemde Derin Takviyeli Öğrenme (DRL) yöntemiyle bir enerji yönetimi yapılmıştır. Bu model Toyota Mirai aracının parametreleriyle oluşturulmuştur. Başta geleneksel araç dinamiği formülleri açıklanmış, ardından elektriksel devre formülleri ele alınmıştır. Yakıt hücresi için yapılan yaklaşımlar tanımlanmış, uygulanan DRL için parametreler seçilmiş ve bu parametreler incelenmiştir. Ödüllendirme için bir maliyet fonksiyonu oluşturulmuştur. Bu maliyet fonksiyonuna girmesi gereken parametreler belirlenmiştir. Bu parametrelerin öncelik sırasına göre ödülleri ile DRL etmeni eğitilmiş ve sonrasında simülasyon ortamında doğrulanmıştır. Bu doğrulama işlemi belirlenen sürüş döngüsü ile gerçekleştirilmiştir. Böylelikle sürüş döngüsünü takip edebilen bir etmen elde edilmiş ve simülasyon ortamında hedeflenen enerji yönetimi sistemi kurulmuştur. Yöntemin başarısı simülasyon sonuçlarıyla gösterilmiştir.

Abstract

In this study, an energy management was performed by Deep Reinforcement Learning (DRL) method in the system by using a vehicle model with a fuel cell stack and battery. This model was created with the parameters of Toyota Mirai car. First, traditional vehicle dynamics formulas are explained, then electrical circuit formulas are discussed. The approaches made for the fuel cell were defined, the parameters for the applied DRL were selected and these parameters were examined. A cost function has been created for rewarding. The parameters that must enter into this cost function are determined. In order of precedence of these parameters, the DRL agent with the rewards was trained and then validated in simulation environment. This method has been validated with selected driving cycle. Thus, an agent that can track the driving cycle was obtained and the targeted energy management system was installed in the simulation environment. Success of applied method is shown in simulation results.

1. Giriş

Günümüzde otomotiv endüstrisi çoğunlukla fosil yakıtlara dayanmaktadır, ancak dünya petrol rezervlerinin güncel tüketim verileriyle ancak 2073 yılına kadar sürdürülebileceği

tahmin edilmektedir [1]. Bu nedenle, otomotiv endüstrisinde karbon ve doğaya zararlı gazların emisyonlarının azaltılması beklenmektedir. Ancak geleneksel sistemlerle bu hedefleri gerçekleştirmek mümkün değildir.

CO₂ salınımını önemli ölçüde azaltmak için mevcut otomobillere bataryalar veya süperkapasitörler gibi ek enerji kaynaklarının eklenmesi bir çözüm olabilir [2].

Son yıllarda yakıt hücreli araçların otomotivde bir çözüm olarak ön plana çıktığı görülmektedir. Hidrojenin enerji kaynağı olarak kullanılıp motordan tekerleğe emisyonuz güç aktarımı olanağı sağlaması ile gündelik sürüş ihtiyaçlarını karşılaması mümkün hale gelmektedir. Yakıt hücrelerinin gecikmeli güç tepkisi nedeniyle, bataryalar veya süperkapasitörlerin kullanılması yakıt hücresinin verimliliğini artırabilir ve rejeneratif frenlemeye imkân sağlar. Bu çözümlerin uygulamaları, Honda FCX Clarity ve Toyota Mirai gibi araçlarda gözlemlenebilir [3].

Bu çalışmada Toyota Mirai aracı ile ilgilenilmiş, bu araç sisteminin, Derin Takviyeli Öğrenme (DRL) teknikleri kullanılarak batarya ve yakıt hücreleri arasında araçtan istenen gücün paylaştırılması hedeflenmiş ve veriler incelenmiştir. Seçilen DRL yöntemi Derin Q Öğrenmesi'dir (DQL). Problem sürekli zamanlı durum uzayı barındırdığı için geleneksel Q Öğrenmesi zayıf kalacaktır. DQL yönteminin doğrulanması US06, UDDS, NEDC ve WLTP sürüş çevrimi için yapılmıştır.

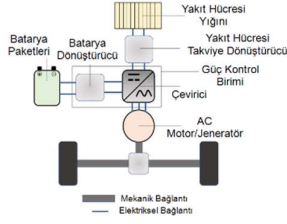
Literatürde Fuel Cell Hybrid Electric Vehicle (FCHEV) araçlardaki enerji yönetim problemi çalışmalarında amaç fonksiyonu içerisinde genellikle 2 adet parametre (SOC ve Hidrojen debisi) kullanılmıştır [4]. Bu çalışmada ise literatürden farklı olarak SOC, yakıt hücresi verimliliği, sistem giriş çıkış güçlerinin oranı ve yakıt hücresi gücü amaç fonksiyonu tanımlanırken ele alınmıştır. Ek olarak, bu çalışmada yakıt hücresi ömrünün uzatılması da hedef alınmıştır. Seçilen yöntem sürekli zamanlı durum uzayında yetersizliğinden dolayı diğer çalışmalardaki gibi geleneksel Q Öğrenmesi değildir. Ancak problemin içinde sürekli zamanlı aksiyon bulunmadığı için literatürde çok popüler olan DDPG (Derin Deterministik Politika Gradyanı) de seçilmemiştir. DQL yöntemi bu problem için yeterli olarak değerlendirilmiştir.

Bu makalenin geri kalanı aşağıdaki gibi yapılandırılmıştır. Araç dinamiği temelleri ve tasarım adımları ikinci bölümde sistem modeli başlığı altında açıklanmıştır. Üçüncü bölümde derin takviyeli öğrenme (RL)'in ana hatlarından bahsedilmiş ve

teorik detayları açıklanmıştır. Dördüncü bölümde simülasyon sonuçları paylaşılmıştır. Beşinci bölümde ise sonuçlar aktarılmış ve geleceğe dair yapılabilecek çalışmalardan söz edilmiştir.

2. Sistem Modeli

Bu makalede ele alınan Toyota Mirai, yakıt hücrelerini temel güç kaynağı olarak görevlendirirken bataryayı ise yardımcı bir kaynak olarak kullanmaktadır [5]. Batarya aynı zamanda araç yavaşlarken veya yokuş aşağı inerken rejeneratif frenleme ile aracın enerji depolamasına yardımcı olmaktadır. Bu çalışmada geriye dönük modelleme tekniği kullanılmıştır, yani model için giriş parametreleri sürüş çevriminden türetilmiştir. Kullanılan modelde çıktı parametreleri hidrojen tüketimi, hidrojen debisi, yakıt hücresi akımı, yakıt hücresi voltajı, batarya voltajı, batarya akımı ve batarya şarj durumu (SOC)'dur. Bu çalışmada sistem direnç kuvvetlerini içeren araç modeli, elektrik motoru, yükseltici-alçaltıcı (DC-DC) dönüştürücüler, batarya, yakıt hücresi ve hidrojen tankları olmak üzere birkaç alt sisteme ayrılmıştır. Sistemin genel yapılandırması Şekil 1'de gösterilmiştir.



Şekil 1. Toyota Mirai FCHEV Şeması [6,7]

2.1. Araç Dinamiği Modeli

Araç dinamiği model bloğu, sürüş çevriminden gelen hız taleplerine dayanarak elektrik motorundan talep edilen tork ve hızı hesaplamayı amaçlamaktadır. Bu blok, aracın x eksenindeki dinamiklerinin denklemlerinden oluşur. İstenen toplam tekerlek torkundan hesaplanan toplam kuvvet, tüm direnç kuvvetlerinin ve ivmelenme nedeniyle oluşan atalet kuvvetinin toplamıdır.

$$\omega_{teker} = \frac{v}{r_{teker}} \quad (1)$$

$$T_{teker} = \left(\frac{1}{2} \cdot \rho \cdot v^2 \cdot A_f \cdot C_d \cdot f_r \cdot M \cdot g + M \cdot g \cdot \sin\theta + M \cdot a_x \right) \cdot r_{wh} + 4 \cdot J_{wh} \cdot \alpha_{teker} \quad (2)$$

$$\alpha_{teker} = \frac{a_x}{r_{teker}} \quad (3)$$

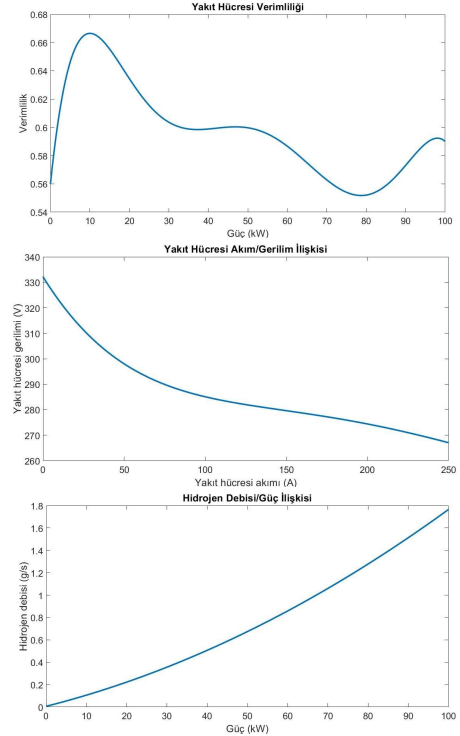
Bu denklemlerde, T_{teker} tekerlek torkunu [Nm], ω_{wh} tekerlek açısal hızını [rad/s], α_{teker} tekerlek açısal ivmelenmesini [rad/s²], g yerçekimi sabitini [m/s²], r_{teker} tekerlek yarıçapını [m], J_{teker} tekerlek dönme ataletini [kgm²], ρ hava yoğunluğunu [kg/m³], v araç hızını [m/s], A_f araç ön alanını [m²], C_d aerodinamik direnç katsayısını, f_r tekerlek sürtünme direnç katsayısını, M ise FCHEV'in tam yük ile toplam kütlelerini [kg] temsil eder. θ yol eğimini [derece] ve a_x uzunlamasına ivmeyi [m/s²] ifade eder. Araç ve güç aktarma organları parametreleri Tablo 1'de listelenmiştir. Yanal dinamiklerin etkisi dikkate alınmamıştır.

Tablo 1. Toyota Mirai FCHEV Araç Parametreleri [8]

Parametreler	Sembol	Birim	Değer
Araç kütlesi	M	kg	1927
Aerodinamik sürtünme katsayısı	C_d	-	0.29
Hava yoğunluğu	ρ	kg / m ³	1.19
Araç ön yüz alanı	A_f	m ²	2.23
Sürtünme kuvveti katsayısı	f_r	-	0.01
Tekerlek yarıçapı	r_{teker}	m	0.316
Tekerlek ataleti	J_{teker}	kgm ²	0.32
Vites oranı	τ	-	9.09
Şanzıman verimliliği	η_{sr}	-	0.98

2.2. Yakıt Hücresi Modeli

Yakıt hücresi sistemi, yakıt hücresi yığını ve hidrojen, hava, su, soğutma ve elektrik devreleri gibi çeşitli elemanlardan oluşmaktadır [9]. Bu sistem, karmaşık iç dinamiklere sahip olduğu için çok fazla hesaplama süresine ihtiyaç duymaktadır [10]. Yakıt hücresi modelinin karmaşıklığı nedeniyle bu çalışmada deneysel sonuçlardan elde edilen verilere dayalı verimlilik grafiği kullanılmıştır. Argonne Ulusal Laboratuvarı tarafından kamuoyuyla paylaşılan veriler yakıt hücresi modelini oluşturmak için kullanılmıştır [10]. Hidrojen debisi, yakıt hücresi sistemi verimliliği ve polarizasyon eğrisi deneysel verilere dayanarak oluşturulmuştur. Verimlilik ve hidrojen debisinin yakıt hücresi çıkış gücüne bağlılığı, Şekil 2'de gösterilmiştir.



Şekil 2. a) Yakıt hücresi verimliliği/güç ilişkisi, b) Yakıt hücresi akım/gerilim ilişkisi, c) Hidrojen debisi/güç ilişkisi [11]

Yukarıdaki eğrilerin polinom olarak ifadeleri aşağıdaki denklemlerde (4...13) verilmiştir [11].

$$n_{yığm} = k_1 \cdot p^6 + k_2 \cdot p^5 + k_3 \cdot p^4 + k_4 \cdot p^3 + k_5 \cdot p^2 + k_6 \cdot p + k_7 \quad (4)$$

$$k_1 = -3.322 \cdot 10^{-11}; k_2 = 1.08 \cdot 10^{-8}; \quad (5)$$

$$k_3 = -1.343 \cdot 10^{-6}; k_4 = 7.986 \cdot 10^{-5}; \quad (6)$$

$$k_5 = -2.3 \cdot 10^{-3}; k_6 = 2.69 \cdot 10^{-2}; k_7 = 0.56 \quad (7)$$

$$n_{sistem} = a_1 \cdot p^5 + a_2 \cdot p^4 + a_3 \cdot p^3 + a_4 \cdot p^2 + a_5 \cdot p + a_6$$

$$a_1 = 4.94 \cdot 10^{-10}; a_2 = -1.357 \cdot 10^{-7}; \quad (8)$$

$$a_3 = 1.335 \cdot 10^{-5}; a_4 = 5.658 \cdot 10^{-4}; a_5 = 7.05 \cdot 10^{-3};$$

$$a_6 = 0.6092$$

$$m_{H_2} = b_1 \cdot p^2 + b_2 \cdot p + b_3 \quad (9)$$

$$b_1 = 8.5 \cdot 10^{-5}; b_2 = 9.1 \cdot 10^{-3}; b_3 = 0.0064 \quad (10)$$

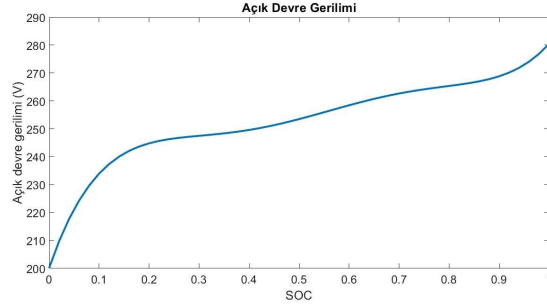
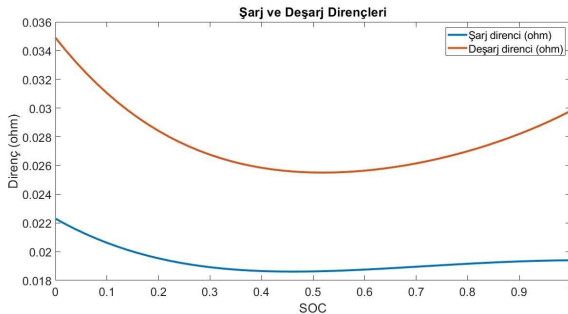
$$V_{yığm} = c_1 \cdot I^5 + c_2 \cdot I^4 + c_3 \cdot I^3 + c_4 \cdot I^2 + c_5 \cdot I + c_6 \quad (11)$$

$$c_1 = -4.587 \cdot 10^{-11}; c_2 = 6.67 \cdot 10^{-8} \quad (12)$$

$$c_3 = -3.553 \cdot 10^{-5}; c_4 = 0.0085; c_5 = -1.028; c_6 = 332.2 \quad (13)$$

2.3. Batarya Modeli

Yapım yılı 2014 olan Toyota Mirai ek bir güç kaynağı olarak 6.5 Ah kapasiteli bir nikel metal hibrit batarya paketi kullanılmaktadır [8]. Literatürde bataryalar Thevenin eşdeğer devre modeli kullanılarak modellenmektedir [12,13]. Mirai ile birbirlerine en benzer batarya verilerine sahip olan, yapım yılı 2004 olan Toyota Prius'tan elde edilen veriler, Idaho Ulusal Laboratuvarı (INL) tarafından toplanmış ve kamuya paylaşılmıştır [14]. Bu modelde ideal bir voltaj kaynağı, şarj ve deşarj dirençleri olmak üzere üç bileşen bulunmaktadır. Bataryanın açık devre gerilimi (OCV) ideal voltaj kaynağı ile temsil edilir ve bataryanın şarj durumuna (SOC) bağlıdır (Şekil 3a). Terminal voltajı, bataryanın dahili şarj ve deşarj dirençleri sırasıyla V_{term} , R_{ch} ve R_{dch} olarak temsil edilir ve bunlar da batarya şarj durumuna (SOC) bağlı parametrelerdir. Şekil 3b ve 3c, bu parametreler ve SOC arasındaki ilişkiyi göstermektedir. Bu çalışmada, sıcaklığın etkisi göz ardı edilmiş ve sıcaklık sürüş çevrimleri boyunca sabit kabul edilmiştir. Şarj ve deşarj akımları sırasıyla I_{ch} ve I_{dch} olarak adlandırılır. Hesaplanma şekilleri denklem (14...19) ile verilmiştir.



Şekil 3. a) Şarj/deşarj dirençlerinin şarj durumuna bağlı grafiği
b) Açık devre geriliminin şarj durumuna bağlı grafiği [14]

Yukarıdaki eğrilerin polinom olarak ifadeleri aşağıdaki denklemlerde (14...19) verilmiştir [14].

$$OCV = m_1 \cdot SOC^5 + m_2 \cdot SOC^4 + m_3 \cdot SOC^3 + m_4 \cdot SOC^2 + m_5 \cdot SOC + m_6 \quad (14)$$

$$m_1 = 1722.4; m_2 = -4747.5; m_3 = 4891.9; \quad (15)$$

$$m_4 = -2312.1; m_5 = 525.72; m_6 = 200.06$$

$$R_{şarj} = n_1 \cdot SOC^4 + n_2 \cdot SOC^3 + n_3 \cdot SOC^2 + n_4 \cdot SOC + n_5 \quad (16)$$

$$n_1 = 0.0056; n_2 = -0.0254; n_3 = 0.0372; \quad (17)$$

$$n_4 = -0.0203; n_5 = 0.0223$$

$$R_{deşarj} = q_1 \cdot SOC^4 + q_2 \cdot SOC^3 + q_3 \cdot SOC^2 + q_4 \cdot SOC + q_5 \quad (18)$$

$$q_1 = 0.0188; q_2 = -0.0547; q_3 = 0.0765; \quad (19)$$

$$q_4 = -0.0457; q_5 = 0.0349$$

2.4. DC/DC Dönüştürücü Modeli

FCHEV bir aracı için yakıt hücresi gerilimini yönetmek için bir dönüştürücü devre entegre edilmelidir. Toyota Mirai'nin mimarisi de bu şekilde tasarlanmıştır. Bunun sebebi alternatif akım senkron motorunun gereksinimi olan voltaj yaklaşık olarak 650 V olmasına rağmen yakıt hücresi yığınının gerilimi 245 V olmasıdır [15]. Bu çalışmada kullanılan modelde bu dönüştürücünün verimliliği 95% olarak ele alınmıştır.

2.5. Güç Paylaşım Fonksiyonu

Talep edilen güç ve yakıt hücresinin anlık çalışma oranına göre, elemanların fiziksel kısıtları göz önünde bulundurularak, bir güç paylaşım fonksiyonu kullanılmaktadır. Talep edilen gücün pozitif olduğu durumda, toplam istenilen güç bataryanın verebileceği azami güç ve yakıt hücresinin anlık sağladığı güç ile karşılanıyorsa (şarj durumunun %0 olması dahil), aradaki fark giriş çıkış güçleri arasında uyumsuzluk olarak okunmaktadır. Talep edilen gücün negatif olduğu ve yakıt hücresinden sağlanan güç ile toplam bataryanın şarj olma kapasitesinden fazla (şarj durumunun %100 olması dahil) olduğu durumlarda ise aradaki fark koruma elemanlarında harcanıyor olarak kabul edilmiştir. Koruma elemanları, mekanik fren veya elektrikli fren direnci olarak düşünülebilir. Bu durumda yakıt hücresinin sağladığı güç koruma elemanlarınca harcandığından kayıp güç olarak okunmaktadır.

Her bir elemana giden güç aşağıdaki denklemlerle ifade edilebilir.

$$P_{FC_{çıkış}} = \max(0, (\min(P_{FC_{maks}}, P_{FC})) \quad (20)$$

$$P_{BAT_{çıkış}} = \max(-P_{BAT_{maks}}, \min(P_{BAT_{maks}}, (P_{talep} - P_{FC_{çıkış}}))) \quad (21)$$

$$P_{koruma} = \min(0, P_{talep} - P_{BAT_{çıkış}} - P_{FC_{çıkış}}) \quad (22)$$

$$P_{FC_{maks}} = 90kW; P_{BAT_{maks}} = 26.5kW$$

3. Derin Takviyeli Öğrenme ile Enerji Yönetim Stratejisi

3.1. Problem Formülasyonu ve Amaç Fonksiyonu

Tasarlanan kontrolcünün sağlaması gereken temel koşullar; performans, güvenlik ve verimlilik olarak ayrılabilir. Talep edilen gücün karşılanma oranı performans, bataryanın yük durumunun belirlenmiş bir aralıkta tutulması güvenlik, yakıt hücrelerinin çıkış gücünün giriş gücüne oranı ise verimlilik olarak ele alınmıştır. Temel koşullara ek olarak, yaşanmayı önlemek için güç değişiminin en aza indirgenmesi de problem tanımına eklenmiştir. Bu seçimler doğrultusunda amaç fonksiyonu aşağıdaki gibi belirtilebilir.

$$\varphi = \max \sum -k_1 |SOC - SOC_{ref}| + k_2 (\eta_{FC}) + k_3 \left(\frac{P_{çıkış}}{P_{talep}} \right) - k_4 |P_{FC}| \quad (23)$$

3. terim $P_{talep} = 0$ olduğu durumda 1 olarak kabul edilmiştir.

3.2. Derin Q-Öğrenmesi

Sutton (2018), Q-öğrenmesi olarak bilinen politikasız bir kontrol algoritmasının keşfinin, takviyeli öğrenmede ilk önemli atılımlardan biri olduğunu belirtmektedir [16].

DQL algoritması derin sinir ağlarının kullanımı sayesinde yüksek boyutlu ve sürekli durum uzayına sahip sistemlerle başa çıkabilmektedir.

Q-öğrenmesinin temelini Bellman optimalite denklemleri oluşturmaktadır. Maksimum kümülatif ödüle ulaşmak Q değerinin Bellman denklemlerinde maksimize edilmesiyle gerçekleşmektedir.

$$Q(s_t, a_t) = E[r_{t+1} + \gamma r_{t+2} + \dots](s_t, a_t) \quad (24)$$

$$Q^*(s_t, a_t) = r_{t+1} + \gamma \max_a Q(s', a) \quad (25)$$

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s', a) - Q(s_t, a_t)] \quad (26)$$

$$\pi^*(s_t) = \operatorname{argmax}_a Q^*(s_t, a_t) \quad (27)$$

Yukarıdaki denklemlerde s, a, t, E, γ , Q^* ve π sırasıyla şunlardır: Durum, aksiyon, güncel zaman, beklenti, indirim faktörü (discount factor), optimal Q fonksiyonu ve politika (policy).

DQL için Q değerinin güncellenmesi aşağıdaki yapıya dönüşmektedir.

$$Q(s_t, a_t; \theta) \approx Q(s_t, a_t) \quad (28)$$

Buradaki “yaklaşık” değer DQL’in yaklaşık hesaplama (function approximation) yapmasından kaynaklanmaktadır. θ sembolü derin sinir ağını temsil eden bir parametredir. Yitim

fonksiyonu (loss) olarak MSE (Mean Squared Error) seçilmiştir.

$$L = \frac{1}{M} \sum (Q_{hedef_i} - Q(s_i, a_i; \theta))^2 \quad (29)$$

Denklem (29)’daki M değeri öğrenme için kullanılan deneyim sayısını ifade etmektedir. Her bir deneyim için bir mutlak hedef Q değeri hesaplanıp öngörülen Q değeri sayesinde hata miktarı hesaplanmaktadır. Deneyim terimi 3.2.2’de açıklanmıştır.

3.2.1. Gözlem Uzayı ve Aksiyon Kütüphanesi

Yakıt hücreli hibrit elektrik aracın enerji yönetim probleminin çözümü için aşağıdaki gözlem uzayı seçilmiştir.

$$S = \{P_{FC_{Norm}}, SOC, P_{talep_{Norm}}\} \quad (30)$$

Sırasıyla yukarıdaki (30) denklemdeki terimler şunları ifade etmektedir: Yakıt hücresi güç çıkışı, bataryanın yük durumu, talep edilen güç. Güç değişkenleri yakıt hücresinin verebileceği maksimum güç olan 90 kW ile bölünerek normalize edilmiştir. Gözlem uzayının değer kümeleri sırasıyla [0, 1], [0, 1] ve [-1, 1]’dir.

Bu değişkenler seçilirken yakıt hücresinin ve bataryanın ömrünü arttırmak ve degradasyonlarını azaltmak hedeflenmiştir [17]. Ayrıca seçilen parametreler arasında minimum bağımlılık olmasına dikkat edilmiştir [18].

Aksiyon kütüphanesi, yakıt hücresinin çıkış gücünün azami gücüne oranını belirten yakıt hücresi yüzdesi değişkenine ekleme ve çıkarmalardan oluşmaktadır. Örneğin yakıt hücresi %30 kapasiteyle çalışıyor ve son seçilen aksiyon -0.01 ise, bir sonraki yakıt hücresi yüzdesi %29 olacaktır.

$$A = \{-0.01, -0.005, 0, 0.005, 0.01\} \quad (31)$$

3.2.2. Ödül Tasarımı

Ödül tasarımında, problem tanımında belirtilen performans, güvenlik ve verimlilik koşulları önceliklendirilmiştir. Ödül fonksiyonu 5 terimden oluşmaktadır. Denklemler (32)-(36) bu terimleri, denklem (37) ağırlıkları, denklem (38) ise ödül fonksiyonunu belirtmektedir. Tüm güç terimleri azami yakıt hücresi çıkış gücü olan 90 kW ile normalize edilmiştir.

$$R_1 = |P_{talep_N} - P_{FC_N} - P_{Bat_N} - P_{koruma_N}| \quad (32)$$

$$R_2 = (P_{koruma_N} < 0)(P_{FC_N}) \quad (33)$$

$$R_3 = |SOC - SOC_{ref}| \quad (34)$$

$$R_4 = |P_{FC}| \quad (35)$$

$$R_5 = (P_{FC_N} > 0)|\eta_{FC_{maks}} - \eta_{FC}| \quad (36)$$

$$W = [k_1 \ k_2 \ k_3 \ k_4 \ k_5] = [1 \ 1 \ 3 \ 0.5 \ 0.0001] \quad (37)$$

$$R = -k_1 R_1 - k_2 R_2 - k_3 R_3 - k_4 R_4 - k_5 R_5 \quad (38)$$

İlk terim, üretilen ve tüketilen güç toplamlarının talep edilen güçle farkıdır. İkinci terim koruma elemanları aktif olduğu durumda yakıt hücresinin harcadığı güçtür. Üçüncü terim şarj durumu hatasıdır. Dördüncü terim işaret bağımsız olarak yakıt hücresinin çıkış gücünün değişimidir. Beşinci terim yakıt hücresi açıkken azami verimlilikten uzaklığıdır.

Buradaki SOC_{ref} ve $\eta_{FC_{maks}}$ değerleri sırasıyla 0.5, ve 0.65’tir. Terimlerin ağırlıkları olay sıklığı ve sinyal skalasına göre seçilmiştir.

Literatürde seyrek ödül (sparse reward) olarak bilinen ortamın etmene “nadiren” ödül vermesi durumudur. Bu, etmenin öğrenme stabilitesini olumsuz etkilemektedir [19]. Bu

sebeple, ödül tasarımı etmene verilen ödülün sürekliliği gözletmiştir.

3.2.2. Uygulanan Geliştirme Yöntemleri

Q Öğrenmesi stabilize problemle bilinmektedir [16]. Bunun sebebi tek bir ağın aynı anda hem hedef değeri hem de öngörü değerini hesaplamasıdır. Buna çözüm olarak iki yöntem geliştirilmiştir. İlki hedef ağ yöntemidir. Politika ağı belli bir periyottan sonra tüm ağırlıklarını hedef ağı kopyalar ve hedefe yakınsama olanağını artırır. İkinci yöntem ise olan öncelikli deneyim tekrarı (ÖDT) kullanmaktır. Deneyimlere öncelik parametresi verilerek yitim fonksiyonunu en çok azaltan deneyimler hafızadan daha sık çekilir [20].

3.3. Öğrenme Süreci

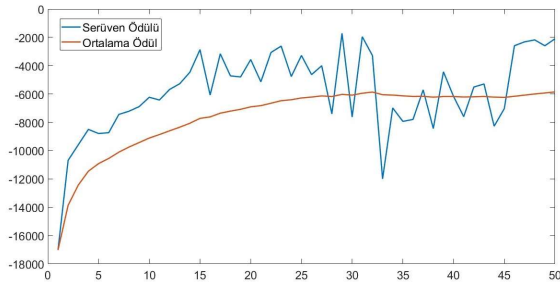
Eğitim sürecinde etmeni mümkün olduğunca farklı durumlara karşı karşıya bırakmak için, popüler sürüş çevrimlerinden olan US06, UDDS, NEDC ve WLTP arka arkaya eklenerek tek bir sürüş çevrimi elde edilmiştir. Doğrulama aşamasında da aynı çevrim kullanılmıştır. Her bir serüven başlangıcında deneyim hafızası sıfırlanmış ve başlangıç şarj durumu %20 ile %80 arasında rastgele bir değerle güncellenmiştir.

Tablo 2’de eğitim boyunca kullanılan sabit parametreler verilmiştir. İndirim faktörünün yüksek seçilmesinin eğitimde başarıyı önemli oranda arttırdığı gözlenmiştir.

Tablo 2: Sabitlenmiş Hiperparametreler

Parametre	Değer
Saklı katman sayısı	3
Nöron dizilimi	15 / 15 / 15
Etkilenim (activation)	ReLU
Yitim fonksiyonu (Loss)	MSE (Mean Squared Error)
İyileştirici (optimizer)	Adam
Öğrenme hızı	0.002
İndirim faktörü	0.9999
Serüven sayısı (Episode)	50
Simülasyon süresi	4897.6s
Örnekleme zamanı	0.1s
Deneyim hafızası (Replay buffer)	1000
Mini-batch boyutu	128

Etmenin öğrenme sürecindeki performansı Şekil 4’te serüven başına alınan ödül ve ortalama ödül ile gösterilmiştir.

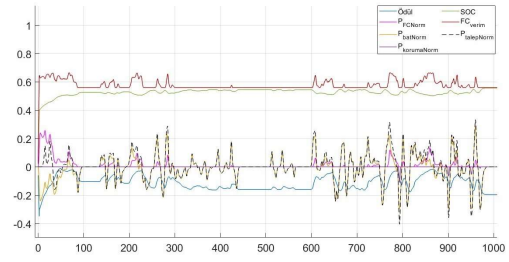


Şekil 4. Etmenin eğitim performansı

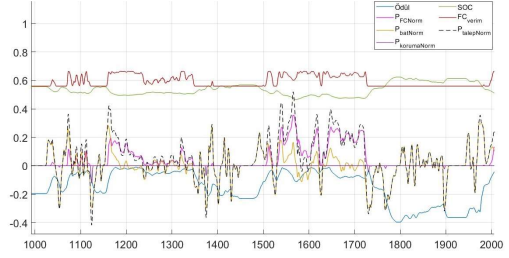
Şekil 4’e göre eğitimin yaklaşık 30. serüvenden sonra etmenin gelişimine katkısının azaldığı görülebilir. En yüksek ödülü 29. serüven sırasındaki etmen kazanmıştır.

4. Simülasyon Sonuçları

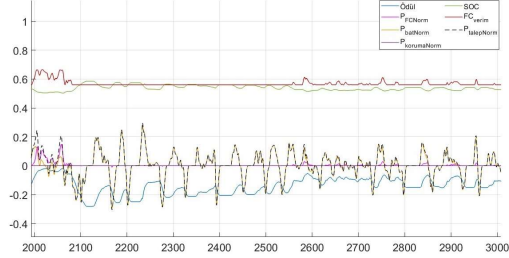
Yaratılan sürüş çevriminin güç talep profili ve etmenin aksiyonlarına göre çevrenin cevapları Şekil 5, 6, 7, 8 ve 9’da görülmektedir. Bu doğrulama koşusunda şarj durumu %40 ilk değerle başlamıştır. Şekil 1’de ilk dakika içinde şarj durumunu referansa getirmek için güç talebi olmasa dahi yakıt hücresinin açıldığı gözlemlenebilir. Yalnızca 3600-3700. saniyeler arasındaki ani ve yüksek güç talepleri tam karşılanamamıştır. Simülasyon boyunca şarj durumunun referans değeri yakınlarında seyrettiği gözlemlenmektedir. İhtiyaç olmadıkça yakıt hücresinin kullanılmadığı görülebilir. Yakıt hücresinin debi değişimini en aza indirmek için, aksiyon uzayındaki değerlerin küçük tutulmasının, ödül üzerinde iyileştirme yapmaktan daha tutarlı bir çözüm olduğu gözlenmiştir.



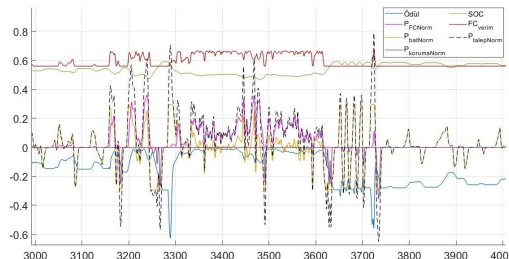
Şekil 5. Doğrulama koşusu, 0-1000 s



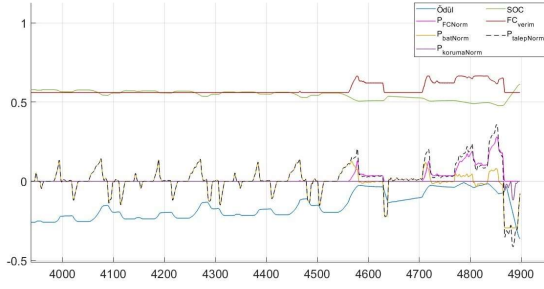
Şekil 6. Doğrulama koşusu, 1000-2000 s



Şekil 7. Doğrulama koşusu, 2000-3000 s



Şekil 8. Doğrulama koşusu, 3000-4000 s



Şekil 9. Doğrulama koşusu, 4000-4900 s

Etmenin batarya ve yakıt hücresi arasında yaptığı seçim, ödül ağırlıklarına hassasiyet göstermektedir. Şarj durumunun istenilen değerden uzaklığına verilen ceza belirli bir değerin altında olduğu durumlarda etmenin yakıt hücresini neredeyse hiç kullanmadığı gözlemlenmiştir. Cezada bu terimin ağırlığını arttırmanın, şarjın düşük olduğu durumlarda, etmeni yakıt hücresinin yüzdesini arttırmaya teşvik ettiği gözlenmiştir.

5. Sonuçlar ve Geleceğe Yönelik Çalışmalar

Yapılan çalışma sonucunda bir yakıt hücreli aracın enerji yönetimi stratejisi için Takviyeli Öğrenme'nin, özel olarak da Derin Q Öğrenmesi'nin başarıyla uygulanabildiği gösterilmiştir. DRL etmeni, parametrelerin önceliklendirilmesine göre belirlenen ödüllerle bir simülasyon ortamında eğitilmiş ve doğrulanmıştır. Doğrulama için US06, UDSS, NEDC ve WLTP çevrimlerinin birleşimi kullanılmıştır. Böylece sürüş döngüsünü takip edebilen bir etmen elde edilmiş ve hedeflenen enerji yönetim sistemi simülasyon ortamında gerçekleştirilmiştir. Yöntemin başarısı simülasyon sonuçları ile gösterilmiştir.

İlerleyen süreçte, aynı probleme DDPG, PPO, SAC gibi diğer Takviyeli Öğrenme yöntemleri uygulanacaktır. Sinir ağı yapısı olarak LSTM'in aynı problemdeki performansı incelenecektir. Daha geniş bir ödül fonksiyonu yelpazesinde etmenin performansı incelenecektir. Katman sayısı, öğrenme hızı, öncelikli deneyim tekrarı hafızası uzunluğu gibi meta-parametrelerin performansa etkisi incelenecektir.

Kaynakça

- [1] J. P. Gurney, "BP Statistical Review of World Energy," Journal of Policy Analysis and Management, vol. 4, no. 2, p. 283, Jan. 1985, doi: 10.2307/3324639.
- [2] L. Castellazzi, S. Ruzimov, A. Bonfitto, A. Tonoli, and N. Amati, "A Method for Battery Sizing in Parallel P4 Mild Hybrid Electric Vehicles," SAE International Journal of Electrified Vehicles, vol. 11, no. 1, pp. 97–111, Sep. 2021, doi: 10.4271/14-11-01-0008.
- [3] S. C. Davis, S. E. Williams, R. G. Boundy, and S. A. Moore, "2016 Vehicle Technologies Market Report," May 2017. doi: 10.2172/1361368
- [4] J. Morales-Morales, I. Cervantes, and U. Cano-Castillo, "On the Design of Robust Energy Management Strategies for FCHEV," IEEE Transactions on Vehicular Technology, vol. 64, no. 5, pp. 1716–1728, May 2015, doi: 10.1109/tvt.2014.2336214.
- [5] Outline of the Mirai. 2017. Available online: www.toyota-europe.com (accessed on 10 March 2023).

[6] Usmanov, U.; Yuldashev, Q. State of the art of fuel cell technology in automotive industry. Int. J. Universum Tech. Sci. 2022, 5-11, 33–40.

[7] Usmanov, U. Control Strategy Optimization of Toyota Mirai Based Fuel Cell Hybrid Electric Vehicles. Master's Thesis, Politecnico di Torino, Turin, Italy, 2022.

[8] Toyota Mirai Technical Specifications vs. FCHV, November 2014. Available online: <https://mag.toyota.co.uk/toyota-miraitechnical-specifications-vs-fchv-adv/> (accessed on 10 March 2023).

[9] Outline of the Mirai. 2017. Available online: www.toyota-europe.com (accessed on 10 March 2023).

[10] S. Luciani and A. Tonoli, "Control Strategy Assessment for Improving PEM Fuel cell System Efficiency in Fuel cell Hybrid Vehicles," Energies, vol. 15, no. 6, p. 2004, Mar. 2022, doi: 10.3390/en15062004.

[11] D3 2016 Toyota Mirai. Toyota Mirai Argonne National Laboratory Testing Results. 2019. Available online: <https://www.anl.gov/taps/d3-2016-toyota-mirai> (accessed on 10 March 2023).

[12] L. Guzzella and A. Sciarretta, Vehicle Propulsion Systems: Introduction to Modeling and Optimization. 2007. [Online]. Available: <http://ci.nii.ac.jp/ncid/BA73456325>

[13] A. Mukhitdinov, S. Ruzimov, and S. Eshkabilov, "Optimal Control Strategies for CVT of the HEV during a regenerative process," 2006 IEEE Conference on Electric and Hybrid Vehicles, Dec. 2006, doi: 10.1109/icehv.2006.352278.

[14] Idaho National Laboratory. Department of Energy Advanced Vehicle Testing. 2004 Toyota Prius—1052 Hybrid Battery Test Results. Available online: <https://avt.inl.gov/sites/default/files/pdf/hev/batteryprius1052.pdf> (accessed on 10 March 2023).

[15] D. J. Pradeep, M. M. Noel, and N. Arun, "Nonlinear control of a boost converter using a robust regression based reinforcement learning algorithm," Engineering Applications of Artificial Intelligence, vol. 52, pp. 1–9, Jun. 2016, doi: 10.1016/j.engappai.2016.02.007.

[16] R. S. Sutton, F. Bach, and A. G. Barto, Reinforcement Learning: An Introduction. Massachusetts: MIT Press Ltd, 2018.

[17] Y. Zhou, L. Huang, X. Sun, L. H. Li, and J. Lian, "A long-term energy management strategy for fuel cell electric vehicles using reinforcement learning," Fuel Cells, vol. 20, no. 6, pp. 753–761, Oct. 2020, doi: 10.1002/fuce.202000095.

[18] B. Xu et al., "Parametric study on reinforcement learning optimized energy management strategy for a hybrid electric vehicle," Applied Energy, vol. 259, p. 114200, Feb. 2020, doi: 10.1016/j.apenergy.2019.114200.

[19] J. Z. Hare, "Dealing with Sparse Rewards in Reinforcement Learning," arXiv (Cornell University), Oct. 2019, [Online]. Available: <https://arxiv.org/pdf/1910.09281.pdf>

[20] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized Experience replay," arXiv (Cornell University), Jan. 2016, [Online]. Available: <https://arxiv.org/pdf/1511.05952>