

# Derin Q-Öğrenimi ile Acil Durum Araçlarının Kontrolü

## Control of Emergency Vehicles with Deep Q-Learning

*Furkan Güney<sup>1</sup>, Hasan Haydar Yıldız<sup>1</sup>, Hilal Çoruh<sup>1</sup> İlhan Tunç<sup>2</sup>, Mehmet Turan Söylemez<sup>1</sup>,*

<sup>1</sup>Kontrol ve Otomasyon Mühendisliği Bölümü  
İstanbul Teknik Üniversitesi

guneyf18@itu.edu.tr  
yildizh16@itu.edu.tr  
coruh16@itu.edu.tr  
soylemez@itu.edu.tr

<sup>2</sup>Mekatronik Mühendisliği Bölümü  
Bursa Teknik Üniversitesi

ilhan.tunc@btu.edu.tr

## 1. Giriş

### Özetçe

Günümüzde, trafik sorunu herkesi yakından ilgilendirmektedir. Fakat acil durum araçları söz konusu olduğunda, bu konu daha da önem kazanmaktadır. Bu çalışma, acil durumlar için trafik riskini azaltmayı amaçlayan bir araştırmayı ele almaktadır. Araştırmanın temel amacı, Derin Q-Öğrenimi yöntemini kullanarak ambulansların hastaları en hızlı ve en uygun rotayla hastaneye ulaştırmasını sağlamaktır. Aciliyet seviyeleri, trafik yoğunluğu, hasta ve ambulans mesafeleri gibi faktörler durum vektörleriyle modellenmiştir. Derin Q-Öğrenimi algoritması, bu vektörleri kullanarak en uygun aksiyonları seçerek ambulansların hastaları taşınması için en etkili rotaları belirlemektedir. Ödül fonksiyonu, hastaların bekleme sürelerine göre önceliklendirilerek ceza fonksiyonuna dönüştürülmüştür. Çalışmada, Derin Q-Öğrenimi ile oluşturulan ajanın durumu öğrenme sonuçları değerlendirilmiş ve öğrenim süreci başarıyla tamamlanmıştır.

### Abstract

This study focuses on reducing traffic risks for emergency vehicles, particularly ambulances. The primary objective of the research is to enable ambulances to transport patients to the hospital using the fastest and most convenient routes through the implementation of the Deep Q-Learning method. State matrices were developed to model factors such as urgency levels, traffic density, patient and ambulance distances. Leveraging these matrices, the Deep Q-Learning algorithm selects optimal actions and determines efficient routes for ambulance transportation. The reward function was designed to prioritize minimizing patient waiting times and was transformed into a penalty function. In the study, the learning results of the agent's state created with Deep Q-Learning were evaluated and the learning process was successfully completed.

Trafik sıkışıklığı dünyanın her yerinde durmaksızın artmaktadır ve bu durumu işe gidip gelenler insanlar için bir engel haline gelmiştir. Artan nüfus ve kentleşmenin bir sonucu olarak, dünya genelindeki şehirlerde ulaşım talebi sürekli olarak yükselmektedir [1]. Özellikle büyük metropollerde, trafik sıkışıklığı yüzünden zamanın verimsiz kullanılması, hava kirliliği ve iklim değişikliğinin etkileri gibi problemlere maruz kalınmaktadır [2]. Ulaşım talebinin artması araç sayısında hızlı bir artış, araç sayısındaki hızlı artış da trafik sıkışıklığını beraberinde getirmiştir. Trafik sıkışıklığının etkisi acil durum araçlarının çalışmasını da etkiler [3]. Trafik sıkışıklığı problemini çözmek için çeşitli akıllı ulaşım sistemleri geliştirilmiştir.

Akıllı ulaşım sistemlerinin en yaygın kullanım alanı trafik ışık kontrolüdür. Kentsel ağdaki araç sayısı hızla arttığından, trafik ışığı kontrolü için etkili mekanizmaların geliştirilmesi gereklidir. Sinyal kontrolünün amacı, kavşak kapasitesini artırmak, gecikmeleri azaltmak ve aynı zamanda trafik aktörlerinin güvenliğini garanti altına almaktır. Ayrıca, yakıt kullanımını ve emisyonları azaltabilir [4]. Bu sistemlerde önceden belirlenmiş zamanlayıcılar kullanılır ve trafik koşullarına bakılmaksızın her döngüde her yaklaşıma yeşil ışık verilir. Bu, yoğun trafiğin olduğu bölgeler için en uygun çözüm olabilir, ancak düşük trafik yoğunluğuna sahip bölgelerde sıra bekleyen araç olmadığı için bu sıra düzeni çok faydalı olmayabilir [1]. Yol ağı rota planlamasının verimliliğini artırmak için birçok uzman ve akademisyen çeşitli çalışmalar yürütmüştür, Dijkstra algoritması bu araştırmaların kesim noktasıdır. Dijkstra algoritmasının, iki nokta arasında en uygun yolu ararken kendi eksiklikleri vardır, ancak yeri doldurulamaz avantajları da vardır [5]. Derin Q-Öğrenimi kullanarak trafik sinyali kontrol problemlerini çözmek, akıllı ulaşım alanında bir araştırma odak noktasıdır. Araştırmacılar, son zamanlarda Derin Q-Öğrenimi yöntemlerine dayanan çeşitli çözümler önermiştir. Bu çözümler,

akıllı ulaşım problemleri için Derin Q-Öğrenimi yöntemlerine dayanan çeşitli çözümler sunmaktadır [6]. Yapılan çalışmalar, trafik sinyal kontrolünde yapay zekanın gücünü kullanarak trafik akışını optimize etmeyi ve sıkışıklığı azaltmayı amaçlayan işbirlikçi derin Q-Öğrenimi yöntemlerinin potansiyelini araştırmıştır [7]. Derin Q-öğrenimi optimizasyon algoritması, acil durum rota planlamasında etkili bir şekilde kullanılmış ve acil durumlar için verimli rota tasarımındaki potansiyelini göstermiştir [8]. Bu analiz sonucunda, acil durum araçlarının trafik sıkışıklığı problemini çözmek için Derin Q-Öğrenimi tekniklerinin kullanıldığı bir çalışma yürütülmesine karar verilmiştir.

Bu çalışma, acil durum hizmetlerinin verimliliğini artırmak ve acil durum müdahale sürelerini kısaltmak için trafik yönetimi stratejilerini optimize etmeyi hedeflemektedir. Bu şekilde, acil durum araçları olay yerine en kısa sürede ve en hızlı rotaları kullanarak ulaşabilir, acil durum müdahalelerinin etkinliği ve başarı şansı artırılabilir. Derin Q-Öğrenimi ile oluşturulan algoritmada acil durum araçlarının hedef kişi veya konuma en kısa sürede varması hedeflenmiş ve başarılmıştır.

SUMO (Simulation of Urban Mobility) ve TraCI (Traffic Control Interface) kütüphaneleri, trafik simülasyonu ve gerçek zamanlı veri alışverişi için kullanılan güçlü araçlardır. SUMO, rota seçimi, trafik lambası algoritmaları veya araç iletişimini simüle etme gibi çeşitli araştırma konularını incelemeye yardımcı olur [9].

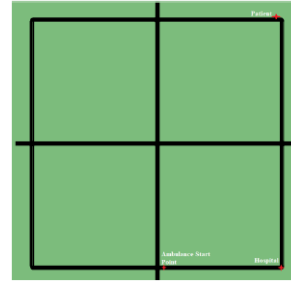
Makalenin 2. başlığı altında hastaların durumu ve sistemin çalışma prensibinden, 3. başlık altında Derin Q-Öğrenimi içerisinde kullanılan işlevsel öğeler ve özelliklerinden, 4. başlıkta yapılan uygulamanın sonuçları ve sonuçların yorumlanmasından, 5. başlıkta ise elde edilen sonuçlar ve geleceğe dönük uygulamalardan bahsedilmiştir.

## 2. Ambulans Aracının Rota Planlaması

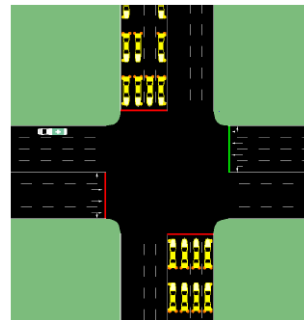
Çalışmanın bu bölümünde Derin Q-Öğrenimi algoritması kullanılarak acil durum araçlarının hastaların aciliyet ve mesafesine göre acil durum aracının kontrolü için optimizasyon yapılmıştır. Acil durum aracı elde olan durum vektörü ile aksiyonlar alıp sonucunda ödül fonksiyonu ile kendini eğitir. Amaç hastaların aciliyeti ve mesafelerine göre hastaları en kısa zamanda hastaneye ulaştırmaktır. Tüm hastaları ulaştırınca bir bölüm biter. Benzetimde üç farklı aciliyet derecesi olan hastalar kullanılmıştır. Hastalar aciliyet seviyelerine göre en acilden en acil olmaya olarak 1, 2 ve 3. derece olarak sınıflandırılmıştır. Acil durum aracının seçim yapabilmesi için tüm hastalar aynı anda benzetimde ortaya çıkmaktadır. Hastalar birbirlerinden farklı noktalarda ortaya çıkmaktadırlar. Aynı ağ üzerinde aksiyon sayısı ve durum vektörü üzerinde değişiklik yapılarak çalışmadan sonuç alınmıştır.

## 3. Acil Araçlar için Derin Q-Öğrenimi ile Rota Planlama

Bu yöntemde, ambulansın her kavşağa geldiğinde sağ, düz veya sol olmak üzere üç aksiyondan birini seçtiği ve bu seçimlerin bir adımı simüle ettiği belirtilmiştir. Durum vektörü 22 elemandan oluşmaktadır ve bu elemanlar hastanın konumu, ambulansın konumu, yol durumu ve hastaların aciliyet seviyeleri gibi bilgileri içermektedir. Model, durum vektörüne uygun bir aksiyon seçmektedir. Ödül fonksiyonu negatif olarak kullanılarak bir ceza fonksiyonu olarak işlev görmektedir. Öğrenme oranı 0.75 olarak belirlenmiştir. Durum vektöründeki her değer, her bölüm (episode) için yeniden oluşturulmakta ve benzetim süresince sürekli güncellenmektedir. Benzetim, 50, 100 ve 200 bölüm olarak güncellenmektedir ve maksimum 8000 adımdan oluşmaktadır. Her adım, programdaki bir benzetim adımını temsil etmektedir. Toplamda 12000 araçlık bir trafik benzetimi yapılmaktadır. Trafik ışıkları sabit bir fazda çalışmakta olup, yeşil ışık süresi 4 saniye ve kırmızı ışık süresi 10 saniyedir. Yol ağı ve acil durum aracının benzetim görüntüleri Şekil 1 ve Şekil 2'de gösterilmektedir.



Şekil 1: Yol Ağı



Şekil 2: Ambulans Görünümü

### 3.1. Durum

Benzetimde kullanılan durum vektörü 22 elemandan oluşmaktadır. Bu elemanlar, acil durum aracının konumu, hastaların konumu, trafik bilgileri ve ağ ağının bilgilerini taşımaktadır. Durum vektörünün elemanları şunları içermektedir:

- *Ajan X pozisyonu, Ajan Y pozisyonu*: Acil durum aracının konumu için x ve y koordinatları

- *Hedef X pozisyonu, Target Y pozisyonu*: Seçilen hastaların konumu için x ve y koordinatları
- *K1 X pozisyonu, K1 Y pozisyonu, K2 X pozisyonu, K2 Y pozisyonu, K3 X pozisyonu, K3 Y pozisyonu*: Komşu kavşakların konumu için x ve y koordinatları
- *K1 araç sayısı, K2 araç sayısı, K3 araç sayısı*: Her bir komşu kavşaktaki araç sayısı
- *K1 trafik ışığı, K2 trafik ışığı, K3 trafik ışığı*: Yeşil durumu 1, kırmızı durumu 0 olarak temsil edilir Her bir komşu kavşaktaki trafik ışığı durumu
- *K1 trafik ışığı döngüsü, K2 trafik ışığı döngüsü, K3 trafik ışığı döngüsü*: Her bir komşu kavşaktaki kalan yeşil ışık süresi
- *Acil hasta sayısı 1, Acil hasta sayısı 2, Acil hasta sayısı 3*: Seviyelerine göre kalan acil hasta sayısı

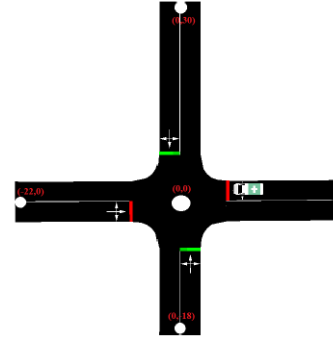
S =	K1 X pozisyonu
	K1 Y pozisyonu
	K2 X pozisyonu
	K2 Y pozisyonu
	K3 X pozisyonu
	K3 Y pozisyonu
	Ajan X pozisyonu
	Ajan Y pozisyonu
	Hedef X pozisyonu
	Hedef Y pozisyonu
	K1 araç sayısı
	K2 araç sayısı
	K3 araç sayısı
	K1 trafik ışığı
	K1 trafik ışığı döngüsü
	K2 trafik ışığı
	K2 trafik ışığı döngüsü
	K3 trafik ışığı
K3 trafik ışığı döngüsü	
Acil hasta sayısı 1	
Acil hasta sayısı 2	
Acil hasta sayısı 3	

Şekil 3: Durum Vektörü

Bu durum vektörü, acil durum aracının aksiyon alması için gerekli bilgileri temsil etmektedir. Bilgiler Şekil 3'te gösterilmiştir. Her bir adımda durum vektörü güncellenmekte ve bu bilgiler kullanılarak uygun bir aksiyon seçilmektedir.

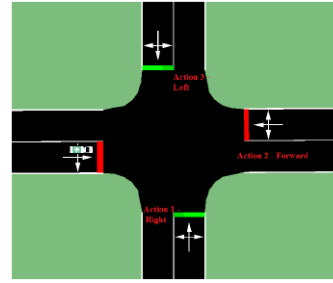
### 3.2. Aksiyon

Aksiyon, acil durum aracının her bir kavşağa geldiğinde bir sonraki hareketi için hangi yolu seçmesi gerektiğidir. Her kavşakta aksiyon alır. Aksiyonu durum vektörüne göre alır. 3 olası aksiyon vardır; sağ, düz ve sol. Bu aksiyonlar ile araç hastaları

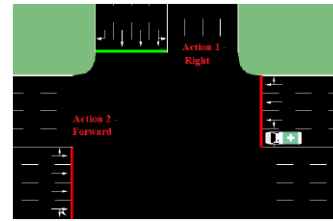


Şekil 4: Konum Koordinatları

bulmaya çalışır. Tüm hastaları bulduğunda 1 bölüm biter. Aksiyon alma aşamasında kenar yola denk gelirse aksiyon sayısı 2 veya 1'e düşer. Şekil 4, Şekil 5 ve Şekil 6'de sırasıyla aksiyon almış aracın nasıl hareket ettiğine yer verilmiştir.



Şekil 5: Aksiyonların Gösterimi



Şekil 6: Aksiyon Seçme Mekanizması

### 3.3. Ödül

Ödül, öğrenme algoritmasında araca geri besleme vererek onun bir sonraki adımda daha iyi sonuç vermesini sağlamaktadır. Bu çalışmada ödül fonksiyonu negatif işaret verilerek ceza fonksiyonu olarak kullanılmıştır. Denklem 1'de ödül algoritmasının hesabının nasıl yapıldığı gösterilmiştir.

- *Anlık adım*: O anki adım sayısı
- *Maksimum adım sayısı*: Benzetimin maksimum adım sayısı, bu projede 8000 olarak ayarlıdır.
- *Acil1*: Aciliyet derecesi 1 olan hasta sayısı

- *Acil2*: Aciliyet derecesi 2 olan hasta sayısı
- *Acil3*: Aciliyet derecesi 3 olan hasta sayısı

Ödül fonksiyonunda temel olarak acil durumdaki hastaların sayıları etkilidir. Fakat bu etki, aciliyet seviyeleri ile orantılı olarak kullanılmak için katsayılar ile çarpılır. Aynı zamanda ödül fonksiyonunun tamamen bağımsız skaler sayılarla ilişkisi yerine aynı zamanda benzetimin ilk ve son adımlarıyla orantılı olması için anlık adım ve maksimum adım sayılarının oranları ile çarpılır. Bu işleme göre ilk adımlarda yapılan hatalara nazaran ilerleyen adımlarda yapılan hatalar optimizasyon için daha etkili olmaktadır. (-) ödül fonksiyonunu ceza fonksiyonuna dönüştürmek için kullanılır. 10 sayısı ise ödül fonksiyonunun çıktılarını belirli bir eşik değerinin üstüne çıkarmak için kullanılır. Ödül fonksiyonu her aksiyon aşamasında ve her benzetim adımında hesaplanır. Bölüm sonunda toplanıp saklanır.

$$Odul = -10 * \frac{AnlikAdim}{Maksimumadim} * (10 * Acil1 + 5 * Acil2 + Acil3) \quad (1)$$

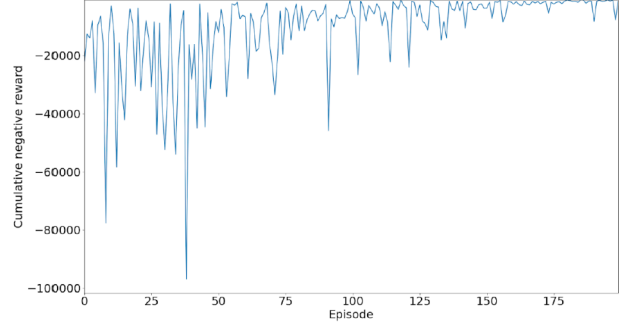
#### 4. Benzetim Sonuçları

Hesaplamalar 2.80 GHz ve 16 GB RAM'e sahip Intel Core i7-1165G7 64 bit makinede yapılmıştır. Toplam 200 bölümden oluşan eğitim süreci 5 saat 50 dakika sürmüştür.

Bu çalışmada, acil durum aracının rota planlaması için Derin Q-Öğrenimi algoritması kullanılmıştır. Ajan, hastaların aciliyet ve mesafelerine göre hastaneye en kısa sürede ulaştırılma amacıyla yönlendirilir. Durum vektörü, acil durum aracının konumu, hastaların konumu, trafik bilgileri ve hastaların aciliyet seviyelerini içerir. Ajan, durum vektörü üzerindeki bilgileri kullanarak uygun aksiyonları seçer ve ödül fonksiyonuyla kendisini eğitir. Benzetim sonuçları, ajanın istenen hedefe ulaştığını ve başarılı bir rota planlaması gerçekleştirdiğini gösterir. Ajanın bölüm sayısı arttıkça daha az ceza alarak hedefe daha iyi ulaşır.

Şekil 7'de yapılan benzetime ait sonuç verilmiştir. Ödül(Reward) olarak geçse de Derin Q-Öğreniminde aslında cezalardan kaçınmak amaçlanmıştır. Dolayısıyla 0 değerine yaklaşmak istenen hedefdir. Yani ajanının, belirlenen sayıdaki bölümleri en az cezayı alarak tamamlamayı hedeflediği söylenebilir. Şekil 7'deki sonuçlara göre, Derin Q-Öğrenimi algoritmasıyla yapılan benzetimlerde istenen sonuca ulaşıldığı ve ajanın öğrenim sürecini başarılı bir şekilde tamamladığı görülüyor. Görselden yola çıkarak ajanın istenen sonuca eriştiği ve öğrenimi başarıyla yaptığı söylenir. Şekil 7'deki sonuçlara dayanarak, ajanın istenen sonuca ulaştığı ve öğrenimi başarılı bir şekilde gerçekleştirdiği söylenebilir. Öğrenme sürecinde ajanın bölüm sayısının uygun bir miktarda tutulması sayesinde daha iyi sonuçlar elde edilebilir. Bu, ajanın öğrenme hızını ve verimliliğini artırabileceği anlamına gelir. Görsel analiz, ajanın performansını değerlendirmek için önemlidir. Sonuçlara bakarak, ajanın hedefini başarıyla gerçekleştirdiği ve doğru kararlar verdiği görülüyor. Bu, Derin Q-Öğrenimi algoritmasının etkili bir şekilde çalıştığını ve ajanın görevini öğrenmek için yeterli veriyi topladığını gösterir.

Sonuç olarak geliştirilen Derin Q-Öğrenimi algoritması ile istenen hedef sağlanmış ve öğrenme başarıyla gerçekleştirilmiştir. Ajan bölüm sayısı ilerledikçe, daha az ceza alarak 0 değerine yaklaşmış ve oldukça iyi sonuçlar ortaya koymuştur.



Şekil 7: 200 bölüm için ödül grafiği

#### 5. Sonuç

Acil durum senaryolarında zamanın kritik bir faktör olduğu düşünüldüğünde, acil durum araçlarının hızlı ve etkin bir şekilde olay yerine ulaşması hayati önem taşır. Bu çalışma acil durum araçlarının rotalarını optimize etmek ve trafikteki akışını yönlendirmek için yenilikçi bir yöntem sunmaktadır. Bu sayede olay yerine varış süresi kısalmış ve acil durum müdahalesi hızlanmıştır.

Bu çalışmada soruna Derin Q-Öğrenimi yöntemi kullanılarak yaklaşmıştır. Deneylerde durum vektörü olarak aciliyet seviyeleri, trafik yoğunluğu, hasta ve ambulans mesafeleri gibi faktörler kullanılmıştır. Ödül fonksiyonu, hastaların bekleme sürelerine göre oranlanmış ve en acil hastaların öncelikli olarak tedavi edilmesi sağlanmıştır. Aksiyon seçenekleri olarak sağ, düz ve sol yolları kullanma tercihi sunulmuştur. Yapılan deneylerde, farklı öğrenme bölümleri kullanılmış ve 200 bölümlük öğrenme en etkileyici sonuçları vermiştir. Ceza fonksiyonunun büyüklüğü her bölüm sonunda azalmış ve grafikler sıfıra yakınsamıştır. Bu sonuçlar, Derin Q-Öğrenimi yönteminin trafik sorununun çözümünde başarılı bir şekilde kullanılabileceğini göstermektedir.

Elde edilen sonuçlar umut vericidir ve gelecekteki çalışmalara yön verebilir. Gelecekteki çalışmalarda, sistemin daha da geliştirilerek birden fazla ambulansın kullanıldığı karmaşık senaryoların modellenilebileceği ve acil durum yönetiminin daha etkin bir şekilde gerçekleştirilebileceği düşünülmektedir. Ayrıca, trafik ışıkları ambulansların geçeceği yollara göre dinamik olarak ayarlanabilir ve ambulansların engelsiz bir şekilde ilerlemesi sağlanabilir. Gerçek hayatta ise ambulansların öncelikli geçiş hakkı sağlayan yöntemlerin kullanılması ve sürücülerin ambulansa yol verme bilincinin artırılması önemlidir.

## 6. Kaynakça

- [1] K. Shingate, K. Jagdale, and Y. Dias, "Adaptive traffic control system using reinforcement learning," *International Journal of Engineering Research and Technology*, vol. 9, 2020.
- [2] İ. Tunç, Ö. Elmas, A. Edem, A. Köroğlu, S. Akmeşe, and M. Söylemez, "Derin Q Öğrenme Tekniği ile Trafik Işık Sinyalizasyonu."
- [3] Vardhana, M., Arunkumar, N., Abdulhay, E. et al. (2019). " Iot based real time traffic control using cloud computing." *Cluster Computing*, 22(Suppl 1), 2495–2504.
- [4] Monireh Abdoos, Nasser Mozayani, Ana Bazzan. "Traffic Light Control in Non-stationary Environments based on Multi Agent Q-learning." *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, October 2011.
- [5] D. Fan and P. Shi, "Improvement of Dijkstra's algorithm and its application in route planning." *2010 seventh international conference on fuzzy systems and knowledge discovery*, Vol. 4, pp. 1901-1904. IEEE.
- [6] G. Han, Q. Zheng, L. Liao, P. Tang, Z. Li, and Y. Zhu, "Deep Reinforcement Learning for Intersection Signal Control Considering Pedestrian Behavior," *Electronics (Basel)*, vol. 11, no. 21, p. 3519, 2022.
- [7] S. Xu et al., "Indoor Emergency Path Planning Based on the Q-Learning Optimization Algorithm," *ISPRS International Journal of Geo-Information*, vol. 11, no. 1, p. 66, 2022.
- [8] S. A. El-Tantawy, H. Abdelgawad, and R. A. Ramadan, "Cooperative Deep Q-Learning for Traffic Signal Control," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 1-16, 2016.
- [9] M. Behrisch, L. Bieker, J. Erdmann, D. Krajzewicz, and C. Rössel, "SUMO – Simulation of Urban Mobility: An Overview," *International Journal on Advances in Systems and Measurements*, vol. 4, no. 34, pp. 308-316, 2011.