

**Ministry of Higher Education and  
Scientific Research  
University of Mosul  
College of Computer Science and  
Mathematics  
Department of Computer Science**



# **Predicting Network Loads for Traffic Violation System Using LSTM and Edge Computing**

**A Thesis Submitted to the Council of the College of  
Computer Science and Mathematics  
University of Mosul  
as a Partial Fulfillment of Requirements  
for the Degree of Doctor of Philosophy in  
Computer Science**

**By**

**Samah Fakhri Aziz Mohammed**

**Supervised by**

**Prof. Dr. Manar Younis Ahmed Jarjees**

---

**2025 A.D.**

**1446 A.H.**

## **Abstract**

The development of the traffic system is one of the challenges that has great interest in most of the countries around the world because they try to enforce traffic laws by implementing efficient traffic violations system. This interest came after the incremented numbers of traffic accidents and their victims and damages of the public and private properties. It is expected with growing numbers of vehicles that the number of the violations will increased. In addition, more IoT devices such as radar and cameras are connected to the traffic violation system. This growth leads to introduce more efficient system with high level of Quality of Service QoS. In such system, latency has to be reduced to meet the requirement of the system according to the capability of the implemented network and devices.

The thesis proposed a traffic violation system that meet the growth of the violation number on the future to build a realble traffic violation system in Mosul city that meet the expected numbers of violations for the next five years. The system design is based on the expected number of violations with considering more devices that are connected to the network using Long-Short-Term Memory (LSTM) deep learning algorithm to predict the expected violations by converting the registered violations to a time series that can be used in prediction. A published dataset of the traffic violation system in Maryland, USA is used in prediction of the violation numbers in the future. The information of the violations is analyzed, it is an undividable piece of data that include (1945992) registered violations during 2022-2024. In order to elevate the performance of the system, the registering of violations is distributed onto edge server according as load balancing. These edge servers are set on clusters initialized by K-mean algorithm considering the number of clusters is equal to 6 clusters. An edge server is set to each cluster centroid as centroid node of the cluster. The system is implemented in Mosul city which managed by the directorate of traffic police. A

python programs and cloud database are used to design and implement a simulation system. A published dataset of violations is used in prediction while Google maps are used to allocate the nodes in the site of duty of the traffic police.

The results showed that implementing LSTM is suitable to predict the time series of the violations in the future with RMSE value equals to 6.4521 at (Lookback = 72 , Batch size = 96 and Epoch number = 30). While, implementing K-Mean clustering algorithm with  $k = 6$  is very efficient because this algorithm produces distributed nodes on each cluster area without outliers. The far nodes are grouped in smaller clusters that converting the disadvantage of K-mean to advantage point. Decentralized service is implemented to distribute the registry of the violations on the edge servers.

To evaluate the performance of the system, two scenarios were implemented to show the differences between the system without and with decentralization. Implementing without decentralization showed incremented latency during incremented requests of (50, 100, 250, 500) clients. According to the available connections equals (50) the latency increased by recognized delay time produced by the devices that the data travels through. While implementing the decentralized system with same number of clients showed that the initial latency is reduced, and latency still at its level till six times the previous implementation. According to the obtained value of RMSE equals (6.4521), the system can manipulate six times of registered violations of the predicted violations. It is worthy to implement prediction and clustering to design a traffic violation system, because it will be scalable in the future.



وزارة التعليم العالي والبحث العلمي  
جامعة الموصل  
كلية علوم الحاسوب والرياضيات  
قسم علوم الحاسوب

# التنبؤ بالأحمال في نظام المخالفات المرورية باستخدام LSTM والحوسبة الحافة

اطروحة مقدمة  
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل  
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في  
علوم الحاسوب

من قبل

سماح فخري عزيز محمد

بإشراف

أ.د. منار يونس احمد جرجيس

## المستخلص

يعد تطوير نظام المرور أحد التحديات التي تحظى باهتمام كبير في معظم دول العالم؛ لأنها تحاول فرض قوانين المرور من خلال تنفيذ نظام فعال لمخالفات المرور، وجاء هذا الاهتمام بعد زيادة أعداد حوادث المرور وضحاياها وأضرار الممتلكات العامة والخاصة، إذ من المتوقع مع زيادة أعداد المركبات أن يزداد عدد المخالفات المرورية، بالإضافة إلى توصيل المزيد من أجهزة إنترنت الأشياء IoT مثل الرادارات والكاميرات بنظام مخالفات المرور، فيؤدي هذا النمو إلى تقديم نظام أكثر كفاءة بمستوى عالٍ من جودة الخدمة QoS، وفي مثل هذا النظام، يجب تقليل زمن الوصول لتلبية متطلبات النظام وفقاً لقدرة الشبكة والأجهزة المنفذة.

تقترح هذه الرسالة نظام مخالفات مرورية يلبي نمو عدد المخالفات في المستقبل لبناء نظام مخالفات مرورية موثوق به في مدينة الموصل، ويلبي الأعداد المتوقعة للمخالفات للسنوات الخمس القادمة، إذ يعتمد تصميم النظام على العدد المتوقع للمخالفات مع مراعاة المزيد من الأجهزة المتصلة بالشبكة باستخدام خوارزمية التعلم العميق للذاكرة طويلة المدى (LSTM) للتنبؤ بالمخالفات المتوقعة عن طريق تحويل المخالفات المسجلة إلى سلسلة زمنية يمكن استخدامها في التنبؤ، تستخدم مجموعة بيانات منشورة لنظام مخالفات المرور في ماريلاند بالولايات المتحدة الأمريكية في التنبؤ بأعداد المخالفات في المستقبل، ثم يتم تحليل معلومات المخالفات، وهي مجموعة بيانات غير قابلة للتجزئة تتضمن (19٤٥٩٩٢) مخالفة مسجلة خلال الفترة ٢٠٢٢-٢٠٢٤، وبغية رفع أداء النظام، يتم توزيع تسجيل المخالفات على خادم الحافة وفقاً لموازنة التحميل، وتحدد خوادم الحافة هذه على مجموعات تم تهيئتها بواسطة خوارزمية K-mean مع الأخذ في الاعتبار أن عدد المجموعات يساوي ٦ مجموعات. تم تعيين خادم حافة لكل مركز كتلة كعقدة مركز الكتلة، وينفذ النظام في مدينة الموصل التي تديرها مديرية شرطة المرور في نينوى، باستخدام برامج بايثون وقاعدة بيانات سحابية لتصميم وتنفيذ نظام محاكاة، يستخدم مجموعة بيانات منشورة للمخالفات في التنبؤ، واستخدم خرائط جوجل لتخصيص العقد في موقع عمل شرطة المرور.

أظهرت النتائج أن تنفيذ LSTM مناسب للتنبؤ بالسلسلة الزمنية للانتهاكات في المستقبل بقيمة  $RMSE=6.4521$  عند ( $Lookback = 72$ ،  $Epoch\ number = 30$ ،  $Batch\ size = 96$ ). في حين أن تنفيذ خوارزمية التجميع K-Mean مع  $k = 6$  فعال للغاية، لأن هذه الخوارزمية تنتج عقداً موزعة على كل منطقة مجموعة بدون قيم متطرفة، تجمع العقد البعيدة في مجموعات أصغر تحول عيب K-mean إلى نقطة مميزة، وتنفذ خدمة لامركزية لتوزيع سجل المخالفات على خوادم الحافة.

جرى تنفيذ سيناريوهين لتقييم أداء النظام لإظهار الاختلافات بين النظام بدون اللامركزية ومعها، أظهر التنفيذ بدون اللامركزية زيادة في زمن الوصول أثناء الطلبات المتزايدة لـ (٥٠، ١٠٠، ٢٥٠، ٥٠٠) عميل، على اعتبار الاتصالات متاحة تساوي (٥٠)، زاد زمن الوصول بمقدار وقت التأخير المعترف به الناتج عن الأجهزة التي تنتقل البيانات من خلالها، بينما أظهر تنفيذ النظام اللامركزي بنفس عدد العملاء أن زمن الوصول الأولي قد انخفض، ولا يزال زمن الوصول عند مستواه حتى ستة أضعاف التنفيذ السابق، وفقاً لقيمة RMSE تساوي (٦,٤٥٢١) وفقاً للقيمة التي تم الحصول عليها، يمكن للنظام معالجة ستة أضعاف الانتهاكات المسجلة للانتهاكات المتوقعة، ومن الأجدر تنفيذ التنبؤ والتجميع لتصميم نظام مخالقات مرورية، لأنه سيكون قابلاً للتطوير في المستقبل.