

A Proposal of a Vehicle License Plates Generation Method

Rodrigo Borges Nogueira¹, Diogo Leonardo Costa¹, Ricardo Staciarini Puttini²

¹*Brazilian Federal Highway Police*

²*University of Brasilia*

Abstract

A new vehicle license plates system has been deployed in some countries of South America. The plates enumeration is based on a common pattern shared by the Mercosul (Southern Common Market) countries, under a formal agreement. However, the proposal signed by these countries follows the same old recipe of distributing a range of plates according to regions and identifying regions/cities/countries explicitly in the plate. We propose a new way of thinking about vehicle license plates, considering the citizen privacy and the effectiveness of OCR (Optical Character Recognition) systems. Every newly generated plate characters' sequence is evaluated against all other existing sequences, in order to avoid confusing characters in a given position of the sequence, which improves the automatic identification of plates. By using plate's images and data collected by the Brazilian Federal Highway Police systems, we could identify the most confusing characters and get statistics on the performance of a real LPR (License Plate Recognition) stations system in operation to motivate the design of our vehicle license plates generation method.

1. Introduction

The Mercosul (Southern Common Market) authorities have decided on a new vehicle license plates system, which has already been deployed in some Latin America countries. In that system, a plate is identified by seven alphanumeric characters, with four letters and three numbers in predefined and fixed positions, according to an agreed distribution of license plate identifiers range for each country **Error! Reference source not found.**

Among the objectives of the Mercosul vehicle license plates system, we highlight, for example, to guarantee the free circulation of vehicles to facilitate productive activities and to enforce the law by fighting illegal drugs trafficking, car robbery and other crimes.

Aiming to increase the citizens' privacy expectations and, at the same time, to improve OCR systems reading performance, we propose the application of a decentralized plate generation method, which outputs a pseudorandom plate, uniformly distributed over the plates domain. The method was originally applied to the Personal Identification Numbers generation problem **Error!**

Reference source not found. and it preserves the capability of identifying the country

that issued any given license plate. This can be accomplished by simply converting a plate p to a corresponding unique number representation $n \in \mathbb{N}^*$ and, then, finding the unique factorization of n . Considering the *Fundamental Theorem of Arithmetic*, the set of distinct prime factors of a number is used to uniquely identify a country or region as the license plate issuer or authority.

Our proposal provides three main properties for a vehicle registration system:

- The issuer of a license plate is always identifiable.
- Any two distinct issuers will be able to generate two disjoint sets of license plates. That is, any given issuer will not generate a license plate that any other issuer generated or might generate in the future. Also, it is not necessary to exchange any message between two different issuers to check whether a plate already exists. To accomplish this, a set of prime numbers is distributed to each license plate issuer on a system setup phase.
- It should be possible to define the domain of all possible license plates of an issuer according to the current number of vehicles in a region/country. Also, the domain is such that it need to consider the estimated number of vehicles in the future. The method additionally accommodates new members that might integrate the unique vehicle license plates agreement afterwards.

The *Brazilian Federal Highway Police (PRF – Polícia Rodoviária Federal)* manages, in different regions of Brazil, a set of cameras for traffic monitoring and automatic car license plate recognition, enforcing traffic laws and preventing accidents and crimes. In order to operate effectively the set of camera, the PRF has deployed a monitoring system to analyze license plates images in real time, dynamically detecting a number of types of irregularities. Three OCR systems from different vendors were deployed in the system and they operate with an average performance of up to 93% when all three OCR tools analyze a given license plate image and agree on the same outcome. However, the overall performance can be relatively low, from 60% to 70%, depending on the weather

conditions, lighting or the license plate physical condition, for instance.

The license plates generation method we propose may improve significantly the performance of the plates' characters automatic reading, whenever the Hamming distance d between two plates is at least two. We define d in terms of the frequency of characters misreading. For instance, we have verified in the PRF database that the characters O, Q and D are the most frequently misread in any given position of a plate string, so in our license plate generation method we avoid the generation of any two plates p_1 and p_2 differing only by these characters in that given position. When plates are generated avoiding confusing characters in a given position, one character misread may allow the identification of the closest existing plate in the space of the Hamming distance d .

Different methods have been proposed in order to increase the matching performance using OCR systems, but not necessarily improving the components or algorithms of an OCR system itself. In **Error! Reference source not found.**, Oliveira-Neto et al proposed a promising strategy to improve the performance of OCR systems by employing a text-mining technique. They considered a scenario with two LPR stations at different locations. Each station could match correctly around 60% of the plates readings. By establishing a correlation between license plates images readings in the two LPR stations, the matching rate improved to around 97%. Their work was developed even further, as in **Error! Reference source not found.**, with a proposition of a self-learning license plate matching algorithm. The basic idea is to develop a model where OCR readings in two different LPR stations are automatically correlated by the system, improving substantially the plates matching rates. In **Error! Reference source not found.**, it is used a (possibly) temporary third LPR station to improve the correlation between the main two LPR stations.

In order to improve a citizen notion of privacy, we further propose that no plate include the region or country of issuance in the plate, neither in the form of characters in the plate identification string, nor as a tag in a specific region of the plate. As OCR technology and systems become more accurate and fast, that piece of information stamped in the vehicle license plate may not be crucial for law enforcement agencies, as it usually require a human to read the region information and most LPR stations does not read it automatically. In our proposal, the region or country can be decoded from a license plate identification string itself, by anyone with a simple application, even when offline or in portable devices, i.e., it is a simple computational operation. The decoding is possible by using the public setup information of the license plates generation method.

On removing the region or country information, a driver and passengers may have the expectation of more privacy, due to the uniformity of all vehicle plates, which is a contribution to an important Mercosul objective of free transit of citizens, for example.

2. Motivation

Regional prejudice and xenophobia are two unfortunate realities in many places around the world. Even inside a country, people discriminate each other in many different ways, as showed in **Error! Reference source not found.**, for instance. Continuous political or economic integrations among countries, like the Mercosul, enable opportunities for the increasing intraregional circulation of people. However, cultural differences or social and economic gaps between countries or regions may continue to create discriminatory patterns.

Many studies have been trying for years to explain prejudice in general, its roots, causes, impacts and possible solutions in psychology, sociology and other areas, such as in **Error! Reference source not found.** and **Error! Reference source not found.** In a great amount of countries, there are regional disparities that create racial and regional stereotypes, as in Brazil **Error! Reference source not found.**, for example.

Regional prejudice is a very complex topic itself and it affects both social sciences and political views. We do not intend to address this problem in details, but solely to make a contribution related to the license plates generation method, making them uniform in all the Mercosul countries. The proposed method may apply whenever the free circulation of vehicles is intended. Our approach relates to how the plate characters are chosen to improve citizens' privacy expectations when driving around in different regions. The generated plates are *visually indistinguishable*, meaning that an ordinary person may not identify, for example, the region from where a vehicle possibly came from when just looking at a plate characters. However, it is easy for a computational device to *decode* such information with little computational power.

Our method has the potential to improve the OCR systems reading performance, by applying rules on the license plate generation to avoid characters misreading. The method intends to add a new possibility to an important research area, regarding LPR installations performance, but not tackling the OCR internals directly. Much research has been conducted on OCR systems for many years, in topics such as image processing and better matching algorithms, as in **Error! Reference source not found.** A 2012 survey describes the research related to the ALPR (*Automatic License Plate Recognition*) state of the art **Error! Reference source not found.**

Most work in this topic cover improvements of license plate extraction from an image, its segmentation and character recognition.

The ALPR has been studied by academy and industry. By 2002, the world's largest LPR installation case at that time was publicly described in **Error! Reference source not found.** They presented the practicalities involved on the setting of 68 lanes in the Port of Los Angeles, showing general aspects of the system, such as physical installations improvements, systems customizations and monitoring strategies. There are many vendors providing OCR solutions for ALPR in the world, mostly using proprietary technology. However, there are also open source solutions, mostly built on top of the *Tesseract OCR Engine*. In **Error! Reference source not found.** is presented a Tesseract study case. *OpenALPR* **Error! Reference source not found.**, which uses Tesseract OCR library and has both free and proprietary distributions, offers OCR and even vehicle color and brand/model identification capabilities.

3. The benefits of a new license plates generation method

The method we propose assumes a central authority or, as in the Mercosul case, an agreement or a treaty among the member states, such that a central committee governs policies and rules as a whole. The license plates are uniform, with indistinguishable design, fonts, colors, shapes, material and other relevant characteristics to make them look identical. In most countries, it is required that a motor vehicle has an official registration and license plates attached to it in the front and in the rear of the vehicle. In other countries/regions or in the case of some vehicle categories, it is required only one plate in the back, such as a motorcycle.

3.1. License plates and privacy

The main objective of indistinguishable license plates in vehicles is that citizens may experience more privacy, blending in with the vehicles of other regions, without get unnecessary or undesirable attention, which is quite common in small cities for example, as well when a plate is clearly from another country.

Even when license plates have shapes, colors and other properties making these plates similar to each other, it may be possible to identify the origin of a vehicle or to notice it out of a group of vehicles, by just reading the sequence of characters. For example, currently, in Brazil, different regions issue license plates according to a pre-distributed domain. Then, given that the Brazilian license plates format is such as AAA-1234, the first three letters are typical in

certain regions. Also, it is usually possible to estimate the age of a vehicle, due to the sequential license plates' issuance.

In the Mercosul case, it was established that each country would have a format based on a fixed position of letters and numbers. For example, in Argentina, all plates have a format of the Figure 1, with a sequence of two letters, three numbers and two letters. Note that the country is clearly exhibited in the design of the license plate.



Figure 1. License plate format in Argentina

As another example, we can cite the German license plates scheme, where the country and region are explicitly coded using the first characters of a plate identification string. Also, there is a sticker identifying the plate issuer. The extensive list of regions in **Error! Reference source not found.** can lead to some confusing scenarios, where there exist plates with a different number of characters and the *entropy* of the plate identification string is relatively low. The later results in more characters necessary to identify a vehicle uniquely. Figure 2 shows a German license plate. The character D in the left-hand side identifies the country, *Deutschland* (Germany), the WOB is the city/region, *Wolfsburg*, and the remaining characters are a random string. The plate string has no more than eight characters.



Figure 2. License plate format in Germany

We propose that license plates have all explicit country or region information removed. Additionally, plates' characters should be generated at random and both letters and numbers are present in any position.

Traditionally, most ALPR systems does not consider the reading of country/region tags. Therefore, the *modus operandi* of police forces have moved significantly to automated solutions. One can argue how the country/region tag is still useful for the law enforcement procedures nowadays and how much damage it might cause if it was removed. This debate lies on balancing citizen's privacy and the use cases for the license plate information we are proposing to remove.

Some restrictions may apply when considering a pure random solution for generating vehicle's license

plates, especially due to requirements such as fighting numerous occurrences of crimes in countries borders. Our solution compensates the removal of the country/region information from the license plates, in such a way that we do not need an online system to search in a vehicles database.

3.2. Improving OCR performance

Based on data provided by the Brazilian Federal Highway Police (PRF, which stands for *Polícia Rodoviária Federal*), we can define an approach to improve OCR performance, in addition to other well-known configurations and procedures, as described in **Error! Reference source not found.** or OCR vendor manuals.

The PRF traffic oversight activities counts on hundreds of cameras spots in different regions of Brazil. The basic setup consists of cameras to capture license plates imagens from the front and the rear of a vehicle. Due to poor light conditions, rainy weather in some regions, dirty license plates and illegible characters, the OCR performance can be as bad as 60% of correctness. The correctness of the OCR cascade was admeasured based on 10000-labeled images, validated by humans, from 5000 vehicles with two license plates each. All images were chosen at random from different camera sets in different locations and time of day.

After the human validation of license plates images, it was possible to verify which are the most confusing characters and its positions in the sequence of characters.

The new Mercosul license plates characters' font is different from the current plates in use in Brazil. The chosen font for the new license plates model is the *FE Engschrift* font, which is designed to make characters' forgery more difficult and to improve OCR systems readings. Even so, the characters that are more frequently misread can be avoided when generating a license plate identification string. Then, for example, if the first character of a newly generated plate is a D, the generator system can check in the existing plates database whether there exists a plate with a character in the first position that is frequently identified incorrectly as a D, such as a 0 or a O.

We can avoid the generation of plates whose characters may be misread due to the existence of another plate with confusing characters in the same position. However, it may not be possible to exclude a big set of license plates, but the plates generator can postpone the generation of plates with confusing characters as long as possible. In addition, it may be helpful to issue license plates with confusing characters to vehicles in different and distant regions. In doing that, it may be possible to infer automatically the characters of a license plate in case of OCR misreading, based on previous readings of

that plate in a given region and time window or by checking it in a central database. This is a complementary strategy to improve LPR stations performance and it can be used along with those proposed in **Error! Reference source not found.**, **Error! Reference source not found.** and **Error! Reference source not found.**, for example.

If an online database with all vehicles license plates is available, we can propose an improvement to automatically identify uniquely a vehicle, even when the LPR station system make an incorrect OCR reading. A possible approach is to filter generated plates identification strings and to use the confusion matrix idea from **Error! Reference source not found.** Each element in the matrix represents the conditional probability $p(r|t)$ of a true character t being recognized as r . On generating a license plate p^{new} , we compare each character of p^{new} with characters in corresponding positions of all existing plates p^{DB} in a central database, which contains all already issued license plates. Consider a character r_i in position $i: 0 < i \leq |p^{\text{new}}|$ of the newly generated license plate. Also, consider a character t_i of any plate p^{DB} from the database. Then, we interpret $p(r_i|t_i)$ as the probability of having the character t_i from p^{DB} being recognized as r_i , which may increase the total probability of a false positive when searching p^{DB} in the database after p^{new} is accepted as a new license plate in the database. Summing up probabilities $p(r_i|t_i)$ for all positions i , we accept a new generated license plate p^{new} in the database whenever $\sum_i p(r_i|t_i) < \mu$, where μ is an arbitrary system parameter.

Furthermore, we can consider that many LPR stations operate with a magnetic loop detector capable of identifying some types of vehicles, such as cars, buses, trucks and motorcycles. Then, with this additional information, we can issue a plate with $\sum_i p(r_i|t_i) > \mu$ to vehicles of different categories.

4. License plates generation method

We propose a method for vehicle plates generation, inspired on the ideas proposed for Personal Identification Numbers in **Error! Reference source not found.** We are using the same method to get a decentralized unique pseudorandom numbers generation, based on some public distributed information of a setup phase.

4.1. Defining a domain in \mathbb{N} for license plates

Let us establish a representation of license plates as natural numbers. Define an injective and invertible function $f: P \rightarrow N$ for any given plate $p \in P$, such that $f(p) = n$, $n \in N \subset \mathbb{N}^*$ and $|P| = |N|$. $|P|$ is the domain of all possible license plates issued by an authority A . Now, we use a

positional base-36 numeral system with symbols (digits) 0 ... 9, A ... Z. The value of $|P|$ depends on the policies defined for the license plates generation system, which is usually a decision of a central authority or a designated workgroup.

We will consider that all symbols can be combined in any position to form a license plate, resulting in $|P| = 36^m$, where m is the number of symbols in a license plate. Besides, each plate is generated by an authority A_i as a unique pseudorandom natural number in a decentralized environment, where $0 < i \leq a$ and a is the number of authorities engaged on a common decentralized vehicle registration system.

4.2. Generating license plates

The intuition for the plates generation is that an authority A_i , $0 < i \leq a$, gets a set Γ_i of prime numbers used to generate plates $p_j^i \in P_i$, represented by natural numbers in base-36 $p_j^{i36} \in N_i \subset \mathbb{N}$, $j \in \mathbb{N}: j \leq |P_i|$. Under the *Fundamental Theorem of Arithmetic*, we select all the prime numbers $\gamma_r^i \in \Gamma_i$, $0 < r \leq |\Gamma_i|$ and natural exponents $\zeta_1^i, \zeta_2^i, \dots, \zeta_{|\Gamma_i|}^i \in \mathbb{N}$ in order to generate the natural number $p_j^{i36} \in N_i$.

$$p_j^{i36} = \prod_{x=1}^{|\Gamma_i|} (\gamma_x^i)^{\zeta_x^i} = (\gamma_1^i)^{\zeta_1^i} \cdot (\gamma_2^i)^{\zeta_2^i} \dots (\gamma_{|\Gamma_i|}^i)^{\zeta_{|\Gamma_i|}^i} \quad (1)$$

For example, for plates of 7 characters, the last plate is, lexographically, $p_j^i = ZZZZZZZ \in P_i$, and $f(ZZZZZZZ) = p_j^{i36} = 78\,364\,164\,095$, for some authority A_i , where the factors of p_j^{i36} are in Γ_i , i.e., $\{5,7,29,197,55987\} \subset \Gamma_i$. The greatest factor of p_j^{i36} defines the authority A_i , that is, no other authority $A_k, 0 < k \leq a, i \neq k$, can issue a plate with the greatest factor equals to 55987. We say that the factor $\gamma_x^i = 55987$ is an element of the *signature set* Γ_i^{sig} , $55987 \in \Gamma_i^{sig}$, of the authority A_i , such that $\Gamma_i^{sig} \subset \Gamma_i$.

In order to generate a random plate p_j^{i36} as a random factored number of equation (1), we can use the efficient algorithm presented in **Error! Reference source not found.** The algorithm generates a number n uniformly distributed over $\{1, \dots, m\}$, where m represents the natural number representation of the last possible plate, considering a lexicographical order.

4.3. Setup Phase

The setup phase consists of distributing the public set of prime numbers Γ_i for each authority A_i . Be

authorities A_i and A_k , $i \neq k$, with expected license plates domains P_i and P_k , respectively. If $|P_i| > |P_k|$, then A_i gets Γ_i capable of generating more plates than A_k .

Reference **Error! Reference source not found.** describes a method to construct the prime numbers set Γ_i for an authority A_i , based on the expected size of $|P_i|$. With a the number of authorities in the system, the method iteratively constructs $\Gamma_1, \Gamma_2, \dots, \Gamma_a$, given $|P_1| > |P_2| > \dots > |P_a|$.

For any two distinct authorities A_i and A_{i+1} , the following holds $|P_i| > |P_{i+1}|$ and $\Gamma_i \subset \Gamma_{i+1}$. Define $\Gamma_{i+1} = (\Gamma_i \cup \Gamma_{i+1}^{sig})$, where Γ_{i+1}^{sig} is a set where all elements (prime numbers) $\gamma_s^{i+1} \in \Gamma_{i+1}^{sig}$ are greater than any element $\gamma_r^i \in \Gamma_i$, where $0 < s \leq |\Gamma_{i+1}^{sig}|$ and $0 < r \leq |\Gamma_i|$, i.e.:

$$\forall \gamma_s^{i+1} \in \Gamma_{i+1}^{sig}, \forall \gamma_r^i \in \Gamma_i: \gamma_s > \gamma_r \quad (2)$$

With sets Γ_i and Γ_{i+1} , authorities A_i and A_{i+1} can generate, at least, $|P_i|$ and $|P_{i+1}|$ plates, respectively. The first prime numbers set Γ_1 for authority A_1 has the smallest prime numbers, $\Gamma_1 = \{2,3,5, \dots\}$. With the smallest primes, A_1 is expected to be the authority who can generate the greatest amount of license plates. Additionally, $\Gamma_1 = \Gamma_1^{sig}$, that is, any generated license plate p_j^{136} , $0 < j \leq |P_1|$, has the greatest prime factor in $\Gamma_1 = \Gamma_1^{sig}$.

In order to generate a plate p_j^{i36} for any authority A_i , it must be included as factors of p_j^{i36} one or more prime numbers $\gamma_r^i \in \Gamma_i^{sig}$, $0 < r < |\Gamma_i^{sig}|$. That way, each prime γ_r^i has the role of a signature for the authority A_i .

4.4. Decoding the vehicle plates authority

Given any license plate p^y issued by an authority A_y , we need to find $y \in \mathbb{N}^*: y \leq a$.

First, we find $p^{y36} = f(p^y)$. Then, factorize p^{y36} and identify the greatest factor γ_{max} . Finally, find the value of y by locating the public set Γ_α^{sig} , such that $\gamma_{max} \in \Gamma_\alpha^{sig}$ and constant $\alpha: 0 < \alpha \leq a$. Then, output $y = \alpha$.

Summing up, given Γ_i for authority A_i , be any license plate in the base-36 representation $p_j^{i36} = \prod_{x=1}^{|\Gamma_i|} (\gamma_x^i)^{\zeta_x^i}$, $0 < j \leq |\Gamma_i|$. Therefore, be Γ_i^{p36} the set of all prime factors of p_j^{i36} . Then, the following relations hold:

$$\Gamma_1 = \Gamma_1^{sig} \quad (3)$$

$$\Gamma_{i+1} = \Gamma_i \cup \Gamma_{i+1}^{sig}, i < a \quad (4)$$

$$\Gamma_i \subset \Gamma_k \text{ when } |P_i| > |P_k| \quad (5)$$

$$\Gamma_i^{\text{sig}} \cap \Gamma_k^{\text{sig}} = \emptyset, \forall i, k \leq a, i \neq k \quad (6)$$

$$\forall p_j^{i_{36}} \in N_i: \max(\Gamma_i^{p_{36}}) \in \Gamma_i^{\text{sig}} \quad (7)$$

$$p_j, p_{j'} \in \bigcup_{k=1}^a P_k$$

with $j, j' \in \mathbb{N}^*$, then $p_j \approx^v p_{j'}$ (8)

We described relations (3), (4) and (5) in section 4.3. (6) means that each authority has its own exclusive signature set. In (7), all base-36 license plates have their greatest factor in the signature set of its own authority. And, in (8), every two plates of any authority are visually indistinguishable from each other.

4.5. Centralized environment

Although the plates' generation system allows a decentralized operation based on a setup information, it can be used in a centralized environment as well, in order to save upfront costs of infrastructure and workers. In this case any plate can still encode an authority and it also can be visually indistinguishable.

5. Conclusions and future work

With a uniform plate system, citizens may experience more privacy when in traffic in regions where they fell some prejudice, for example. The visual indistinguishability is a benefit, even though the authority decoding is an easy process.

We have proposed more control on the vehicle license plates generation, in order to enhance the effectiveness of ALPR systems.

As a future work, we intend to implement a license plate generator as an open source tool.

6. References

[1] Grupo Mercado Común (GMC) (2014), 'Resolución - 033/2014 – Patente e Sistema de Consultas sobre Veículos do Mercosul'; http://www.mercosur.int/innovaportal/v/5758/2/innova.front/resoluciones_2014 (22 August 2016).

[2] Nogueira, R. B. and Puttini, R. S. (2015) 'On Unique Personal Identifiers', *Journal of Internet Technology and Secured Transaction* 4.1, pp. 366-72.

[3] Oliveira-Neto, F. M., Han, L. D. and Jeong, M. (2009) 'Tracking large trucks in real time with license plate recognition and text-mining techniques', *Transportation Research Record: Journal of the Transportation Research Board* 2121, pp. 121-127.

[4] Oliveira-Neto, F. M., Han, L. D. and Jeong, M. K. (2013) 'An online self-learning algorithm for license plate matching', *IEEE Transactions on Intelligent Transportation Systems* 14.4, pp. 1806-1816.

[5] Hargrove, S. R., Lim, H. and Han, L. D. (2016) 'Expanding License Plate Matching Capabilities with Secondary Self-Learning Algorithm', *Transportation Research Record: Journal of the Transportation Research Board* 2594, pp. 51-60.

[6] Frayssinet, F. (2015), 'Latin American Migrants Suffer Prejudice in Their Own Region', *Inter Press Service*; <http://www.ipsnews.net/2015/02/latin-american-migrants-suffer-prejudice-in-their-own-region/> (10 September 2016).

[7] McLeod, S. A. (2008), 'Prejudice and Discrimination'. *SimplyPsychology website*; <http://www.simplypsychology.org/prejudice.html> (10 September 2016).

[8] Pettigrew, T. F. (1959) 'Regional differences in anti-Negro prejudice', *Journal of abnormal psychology*, 59(1), 28.

[9] Batista, J. R., Leite, E. L., Torres, A. R. and Camino, L. (2014) 'Black and northeasterners: similarities in racial and regional stereotypes', *Rev. psicol. polít. vol.14 no. 30* São Paulo, ISSN 1519-549X.

[10] Chang, S., Chen, L., Chung, Y. and Chen, S. (2004) 'Automatic License Plate Recognition', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, No. 1.

[11] Du, S. Ibrahim, M., Shehata, M. and Badawy, W. (2013) 'Automatic License Plate Recognition (ALPR): A State of the Art Review', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol 23, Issue 2.

[12] Elovic, P. and Scheppmann, L. (2002) 'Practicalities of Implementing OCR: A Case Study of the World's Largest Installation', *TOC 2002 Americas*, Miami, Florida.

[13] Patel, C., Patel, A. and Patel, D. (2012) 'Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study', *International Journal of Computer Applications* (0975 – 8887). Vol 55.

[14] OpenALPR Website. <http://www.openalpr.com> (Jan 29, 2017).

[15] Presse- und Öffentlichkeitsarbeit (2016) 'Kfz-Kennzeichen und auslaufende Kennzeichen in Deutschland', Stand: 28.12.2016; http://www.kba.de/SharedDocs/Publikationen/DE/Presse/kfz_kennzeichenliste_faltblatt_pdf.pdf?__blob=publicationFile&v=33 (30 January 2017).

[16] Oliveira-Neto, F. M., Han, L. D. and Jeong, M. K. (2012) 'Online license plate matching procedures using license-plate recognition machines and new weighted edit distance', *Transportation research part C: emerging technologies* 21.1, pp. 306-320.

[17] Shoup, V. (2009) 'A Computational Introduction to Number Theory and Algebra', Cambridge University Press; 2 Edition.

7. Acknowledgements

The authors thank the Technological Solutions team of the Brazilian Federal Highway Police for the support and for allowing us to access all statistical data we needed.