

Traffic Police 2017 (TRAP-2017)

With the increasing amount of traffic information collected through automatic number plate reading systems (NPRS), it is highly desirable for police activities and investigations to find meaningful traffic patterns from the accumulated massive historical dataset in order to identify potential criminal behaviors. Nowadays, NPRS sensors are widely spread on Italian highways making the collection of traffic data more accessible to Italian National Police, *Polizia di Stato*.

However, analyzing traffic data in order to find potential criminal behaviors is challenging due to the huge size of the dataset and the complexity and dynamics of traffic phenomena.

The focus of this conference is to develop and employ automatic traffic analysis systems that can detect, track and in general understand the behavior of road-users in order to identify criminal behaviors. For example, most of the criminal activities on Italian highways are perpetrated by criminal drivers in rest stop or service areas. Those drivers often exhibit a recurring *highway-dart-in-and-out* behavior aimed at moving from one service area to another in order to find a suitable target as part of their victim selection process.

The scientist, or research group (RG), is totally free to present contributions with or without the use of the datasets that are provided below. In case of studies carried out with our dataset, the scientist (or RG) is again totally free to try to satisfy the shared-task goal, or to face any other open research problem.

List of Topics

This conference aims to gather not only *data mining* researchers interested in these topics, but also traffic researchers and decision makers. Original research papers related to the following topics, as well as papers regarding other topics that could be relevant in this area, are welcome:

- detection and tracking of road users (vehicles, bikes, trucks, etc.);
- behavior understanding of road users;
- automatic understanding of the environment in traffic scenarios;
- applications related to traffic surveillance;
- vehicle accident analysis.

Using:

- outliers detection and understanding;
- clustering and conceptual clustering;
- process mining;
- inductive logic programming;
- deep learning and classification.

Dataset

Aiming at focusing the efforts on a common goal, a dataset is provided. It contains transits recorded using several gates. Records are referred to the transits of a limited area of Italy, in which gates are homogeneously distributed. The plates are coherently anonymized (a plate is always referred using the same ID). The dataset can be downloaded filling a little form about the affiliation, and other preliminary non-binding indications.

Shared task

With the intent to unify the way experimental results are evaluated and to push research towards the development of a real working system supporting police activity, a shared task is proposed.

The main practical goal is to identify itineraries that could imply a criminal intent. The scientist (or RG) is free to define the concept of itinerary, formalizing it functionally to the proposing approach. Other open data can be integrated in the itinerary features or not, up to the scientist (or RG). Criminal intents could be described, for instance, as follows:

- the sequential visit of service areas each facing the other;
- the sequential visit of service areas in the same direction;
- transits inconsistent under the space-time point of view, that are the proof of cloned plates;
- combinations of the previous cases involving several plates and possibly the same criminal organization;
- and so on.

There are no restrictions about methods or techniques. There are only two requirements to participate in this shared task, that are:

- the use of a virtual machine (VM) to send us the system, such a strategy will make it easier to validate each participant's results;
- the representation formalism of the output (details can be found below), in this way our systems will be able to build an overall report about results, in order to better evaluate the final outcome.

Final system submission

The system must be executable via the shell script: “mySoftware -i DATASET_PATH -o OUTPUT_PATH”. The developed systems must be provided in a VM having as maximum characteristics: 50 Gb (single disk), 4 cores and 16 Gb of RAM. The submissions will be loaded on a VMware ESXi6 platform. Then, if the scientist (or RG) intends to use VMware, the compressed folder of the VM is expected as submission. Otherwise, the VM must be provided compressing the ovf/ova file. The final VM, as compressed file, must be loaded on a web file manager (e.g. Dropbox, Google Drive, OneDrive, etc.) using an account owned by the scientist (or RG). The link of the VM can be provided us using the form in the side panel. The form will ask the link of the file that we will download directly from your web file manager.

Output representation formalism

The BNF grammar has the form:

- `<pattern_set>` ::= CURLY_OPEN <patterns> CURLY_CLOSE
- `<patterns>` ::= [<pattern> COMMA] <pattern>
- `<pattern>` ::= PAREN_OPEN ID COMMA TIME_SLOT <pattern_seq> PAREN_CLOSE
- `<pattern_seq>` ::= SQUARE_OPEN <pattern_part> SQUARE_CLOSE
- `<pattern_part>` ::= SQUARE_OPEN <segments> SQUARE_CLOSE
- | SQUARE_OPEN <segments> SQUARE_CLOSE <repetitions>
- `<segments>` ::= [<segment> COMMA] <segment>

- <segment> ::= PAREN_OPEN ID COMMA ID COMMA NUMBER PAREN_CLOSE
- <repetitions> ::= REPS NUMBER

An example of the output file "patterns" is:

- {
- (pattern_1, 00_24
- [
- [(point_14, point_15, 223), (point_15, point_16, 310)]^3,
- [(point_16, point_17, 307)]^2, [(point_19, point_20, 1405)],
- [(point_15, point_16, 509)]^2
-]
-),
- (pattern_2, 14_20
- [
- [(point_14, point_15, 323), (point_15, point_16, 210)]^0,
- [(point_16, point_17, 347)]^2, [(point_15, point_16, 510),
- (point_15, point_16, 415)]^2
-]
-),
- }

Lexical matching table:

TOKEN NAME	LEXICAL FORMAT	EXAMPLES
ID	[A-Za-z][A-Za-z0-9_]*	id001, ID_002, n34
NUMBER	[0-9]+	0, 100, 1024
PLATE	@[A-Z0-9][A-Z0-9][A-Z0-9][A-Z0-9][A-Z0-9]	@AA98A, @AB987
TIME_SLOT	[0-2][0-9]_[0-2][0-9]	14_20, 00_24, 22_06
COMMA	,	,
PAREN_OPEN	((
PAREN_CLOSE))
CURLY_OPEN	{	{
CURLY_CLOSE	}	}
SQUARE_OPEN	[[
SQUARE_CLOSE]]
REPS	^	^

This grammar defines the formalism to produce the results to be submitted together with the system. The output will contain one or more patterns representing behaviors as described above or having analogous criminal intents.

During the day there are changing traffic conditions, thus for each pattern there is a time slot of validity. The time slot is represented as "hh_hh". In order to express "it doesn't matter", the scientist (or RG) can use the form "00_24". Note that it is possible to express time slots taking two days, e.g. using "22_07".

A pattern could be composed by repeating sub patterns. In such a case the minimum number of loops are reported. Let us consider for instance the sub pattern: "[(point_x, point_y, 523), (point_y,

point_z, 410)]³". There is a first pair of gates having an average travel time of 5' and 23", whereas the second pair has an average travel time of 4' and 10". After the power symbol there is the number of loops.

In general, the average time has the following notation: "dddhhmmss". Where "d" stands for days, "h" for hours, "m" for minutes and "s" for seconds. Starting from left, the sequence of digits equal to zero can be omitted until a digit other than zero is found.

Submission Guidelines

All papers must be original and not simultaneously submitted to another journal or conference. We require a full paper with a limit of 15 pages.

All the selected papers will be included in the Proceedings of TRAP-2017 and published in Springer Series LNCS/AISC. Then, they must be in English and must be formatted according to the relative guidelines. Sample LaTeX2e and WORD files are available at <http://www.springer.com/it/authors-editors/book-authors-editors/book-manuscript-guidelines>. It is not required to submit a cover page first.

Please register and submit your paper at <https://easychair.org/conferences/?conf=trap2017>.

Committees

Program Committee

- Prof. Dino Pedreschi, University of Pisa
- Dr. Fosca Giannotti, Information Science and Technology Institute of the National Research Council at Pisa
- Prof. Stefano Ferilli, University of Bari
- Prof. Federico Chesani, University of Bologna
- Prof. Claudia Diamantini, Marche Polytechnic University
- Dr. Marco Montali, University of Bolzano
- Dr. Fabiana Lanotte, Digital Transformation Team
- Dr. Giorgia Lodi, AgID
- Ing. Giacinto Occhiogrosso, Italian National Police
- Dr. Tommaso Fornaciari, Italian National Police
- Dr. Fabio Leuzzi, Italian National Police
- Dr. Salvatore Iengo, Italian National Police
- Dr. Onofrio Febraro, Italian National Police
- Dr. Fulvio Rotella, Italian National Police
- Dr. Emiliano Del Signore, Italian National Police

Organizing committee

- Dr. Fabio Leuzzi
- Dr. Onofrio Febraro

Invited Speakers

- Dr. Diego Piacentini, Government Commissioner for the Digital Transformation Team (tbc)

Venue and key dates

The conference will be held in via Tuscolana 1556, 00174, Roma (Italy), October 25-26, 2017.

Key dates:

- shared-task system submission deadline: ~~July 21, 2017~~ **August 4, 2017**;
- paper submission deadline: August 21, 2017;
- paper acceptance notification: September 18, 2017;
- conference registration opening: September 19, 2017;
- camera-ready paper version: October 2, 2017;
- conference registration deadline: October 9, 2017.

Contact

fabio.leuzzi@poliziadistato.it
onofrio.febraro@poliziadistato.it

Sponsors

Sponsored by **Oracle S.p.A.**, that, among its contributions and just for this conference, provides a software¹ that can be broadly used to fulfill the objectives of the shared task.

The best paper award is sponsored by Divitech S.p.A..

¹ <http://www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html>