BibTeX:

Combining a Gauss-Markov model and Gaussian process for traffic prediction in Dublin city center

François Schnitzler
Technion
Fishbach Building
32000 Haifa, Israel
francois@ee.technion.ac.il

Thomas Liebig
TU Dortmund University
Artificial Intelligence Group
Dortmund, Germany
thomas.liebig@tu-dortmund.de

Shie Mannor
Technion
Fishbach Building
32000 Haifa, Israel
shie@ee.technion.ac.il

Katharina Morik
TU Dortmund University
Artificial Intelligence Group
Dortmund, Germany
katharina.morik@tu-dortmund.de

ABSTRACT

We consider a city where induction-based vehicle count sensors are installed at some, but not all street junctions. Each sensor regularly outputs a count and a saturation value. We first use a discrete time *Gauss-Markov model* based on *historical data* to predict the evolution of these saturation values, and then a *Gaussian Process* derived from the *street graph* to extend these predictions to all junctions. We construct this model based on *real data collected in Dublin city.*

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Markov processes, multivariate statistics, stochastic processes, time series analysis; I.2.6 [**Artificial Intelligence**]: Learning—*parameter learning*; J.7 [**Computer in Other Systems**]: Real time

Keywords

traffic prediction, Gaussian Process, Gauss-Markov, autoregressive, smart cities, time series, spatio-temporal

1. INTRODUCTION

In the Greater Dublin Area, 750 (4%) junctions are covered by one or several SCATS (Sydney Co-ordinated Adaptive Traffic System) vehicle count sensors. Our goal is to provide estimates of the saturation at each junction, for the current and future times, whereas our previous work [1] only did so for each junction at the current time.

High traffic saturation (cars/km) co-occurs with low traffic flux (cars/hour) and is an indicator for congestions [3].

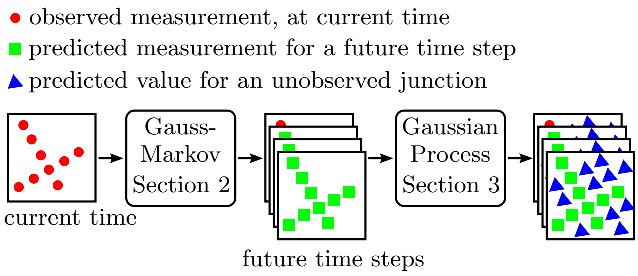

Figure 1: Future measurements are estimated by a Gauss-Markov process (Section 2). Estimates for junctions without sensors, are provided by a Gaussian Process (Section 3).

Our work can be used for online signaling and trip planning.

The urban street network is a graph (V, E), where the vertices V are the junctions and the edges E the street segments. Let u be the set of unobserved junctions, with no SCATS sensor, and $-u = V \backslash u$ the junctions with sensors. The saturation of a junction v_i at a time t is a continuous random variable $y_{i,t}$. Furthermore, $\mathbf{y}_{u,t} \equiv \{y_{i,t}\}_{i:v_i \in u}$.

We combine two components to obtain an estimate of the saturation of all junctions at future time steps, $\mathbf{y}_{V,t+\Delta_t}$, conditioned on the current observations, $\mathbf{y}_{-u,t}$ ($\Delta_t \in \mathbb{N}^0$).

The first one, $P(\mathbf{y}_{-u,t+\Delta_t}|\mathbf{y}_{-u,t})$, models historical measurements. It can estimate future measurements $\hat{\mathbf{y}}_{-u,t+\Delta_t}$, based on the current observations $\hat{\mathbf{y}}_{-u,t}$:

$$\hat{\mathbf{y}}_{-u,t+\Delta_t} = E(\mathbf{y}_{-u,t+\Delta_t}|\hat{\mathbf{y}}_{-u,t}) \ . \tag{1}$$

The second is a Gaussian Process (GP) based on the street network and defining a multivariate Gaussian distribution $P(\mathbf{y}_{V,t})$ over the saturations at all junctions. Conditioning this distribution on $\mathbf{y}_{-u}$ provides $P(\mathbf{y}_{u,t+\Delta_t}|\mathbf{y}_{-u,t+\Delta_t})$ and allows to estimate saturations at junctions without sensors:

$$P(\mathbf{y}_{u,t+\Delta_t}|\mathbf{y}_{-u,t}) \approx P(\mathbf{y}_{u,t+\Delta_t}|\hat{\mathbf{y}}_{-u,t+\Delta_t}) \ . \tag{2}$$

Figure 1 illustrates the resulting prediction procedure.

2. GAUSS-MARKOV

A linear dynamical system models the evolution of a set of state variables $\mathbf{y} \in \mathbb{R}^p$, where we omit the subscript $-u$:

$$\mathbf{y}_{t+1} = A_t \mathbf{y}_t + \mathbf{w}_t \quad (3)$$

$$\mathbf{w}_t \sim \mathcal{N}(\bar{\mathbf{w}}_t, \Sigma_{w_t}) \ . \quad (4)$$

$x_1 \sim \mathcal{N}(\bar{\mathbf{y}}_0, \Sigma_0)$, a multivariate Gaussian distribution of mean $\bar{\mathbf{y}}_0$ and covariance matrix Σ_0. The Kalman filter can compute $P(\mathbf{y}_{t+\Delta_t}|\mathbf{y}_t) = \mathcal{N}(\hat{\mathbf{y}}_{t+\Delta_t}, \hat{\Sigma}_{t+\Delta_t})$ recursively.

Sensor measurements were collected from 2013-01-01 to 2013-05-14[1] by 512 (470 non trivial ones) vehicle count sensors located in central Dublin. We average all measurements received on non-overlapping 4 minutes intervals, because of missing values, and model the resulting averages from 5am to 12am. The parameters A_t, $\bar{\mathbf{w}}_t$, Σ_{w_t} change for every time step but are identical for every day. So are $\bar{\mathbf{y}}_0$ and Σ_0.

Following the methodology of [6], each matrix A_t is learned using (averaged) measurements for $t' \in \{t - \delta_t, \ldots, t + \delta_t\}$, weighted by a Gaussian kernel: $exp(-(t-t')^2/\delta_t)$. We arbitrarily use $\delta_t = 3$. For each matrix A_t, each row $\mathbf{r}_{i,t}$ is estimated using an elastic net [7] and ten-fold cross-validation. Σ_0 and each Σ_{w_t} are diagonal covariance matrices estimated by maximum likelihood. Alternatively, penalized estimation algorithms such as the graphical lasso [2] could be used.

3. GAUSSIAN PROCESS

$P(\mathbf{y}_{u,t+\Delta_t}|\mathbf{y}_{-u,t+\Delta_t})$ is derived from a GP regression framework modeling traffic saturation values of all junctions at a given time, similar to [5]. Multiple sensors at a junction are averaged. For each vertex v_i, we introduce a latent variable f_i, the true traffic saturation at v_i:

$$y_i = f_i + \epsilon_i \quad (5)$$

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2) \ . \quad (6)$$

We assume that the random vector of all latent variables follows a GP: any finite set $\mathbf{f} = \{f_i\}_{i=1,\ldots,M}$ has a multivariate Gaussian distribution. Therefore, the vector of observed traffic saturations ($\mathbf{y}_{-u}$) and unobserved traffic saturations ($\mathbf{d}_u$) follows a Gaussian distribution

$$\begin{bmatrix} \mathbf{y}_{-u} \\ \mathbf{d}_u \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K_{-u,-u} + \sigma^2 I & K_{-u,u} \\ K_{u,-u} & K_{u,u} \end{bmatrix}\right), \quad (7)$$

where I is an identity matrix, K the so-called kernel and $K_{u,-u}$, $K_{-u,-u}$, $K_{u,u}$, and $K_{-u,u}$ the corresponding entries of K. Conditioning on $\mathbf{y}$ produces $P(\mathbf{y}_{u,t+\Delta_t}|\mathbf{y}_{-u,t+\Delta_t})$.

We use the common regularized Laplacian kernel function

$$K = \left[\beta(L + I/\alpha^2)\right]^{-1}, \quad (8)$$

where α and β are hyperparameters. L denotes the combinatorial Laplacian, $L = D - G$. G denotes the adjacency matrix of the graph $\mathcal{G}$ and D a diagonal matrix with entries $d_{i,i} = \sum_j G_{i,j}$. Variables adjacent in $\mathcal{G}$ are highly correlated.

4. DISCUSSION

We have described a combination of two models able to respectively predict future traffic saturations at junctions with sensors and to extend these predictions to junctions without sensors, in a city. To the best of our knowledge, no similar model has been proposed before.

A similar approach was proposed to provide dynamic cost predictions for a trip planner in the same workshop [4]. Instead of a linear dynamical system (LDS), a spatio-temporal Markov random field (STMRF) is used. It models discretized saturation values only, and inference is approximated by belief propagation whereas it is computationally tractable and performed exactly in LDS. Our model also has a finer temporal resolution. Therefore, it can be used for signaling or online adaptation of the route in addition to offline trip planning. Comparing these two models in terms of precision and speed would be interesting.

The Gauss Markov model assumes the dynamics are linear, first-order Markov and perturbed by Gaussian noise. More refined models could be considered and might lead to better estimations.In particular, we could assume the measurements are noisy observations of a hidden process.

Other information could also be leveraged. For example, the street network could be used to derive a prior on the coefficient of the transition matrix, influencing the model only. Irregular, pointwise traffic estimation (for example based on mobile phones or GPS) could be integrated into the Gaussian Process to produce finer saturation estimates. Finally, different dynamics could be estimated and used in the presence or the absence of rain, modifying both the model and the estimation process.

[1] http://dublinked.ie/datastore/datasets/dataset-305.php

5. ACKNOWLEDGMENTS

This work was supported by the European FP7 project INSIGHT under grant 318225.

6. REFERENCES

[1] A. Artikis, M. Weidlich, F. Schnitzler, I. Boutsis, T. Liebig, N. Piatkowski, C. Bockermann, K. Morik, V. Kalogeraki, J. Marecek, A. Gal, S. Mannor, D. Gunopulos, and D. Kinane. Heterogeneous stream processing and crowdsourcing for urban traffic management. In *Proceedings of the 17th International Conference on Extending Database Technology*, page (to appear), 2014.

[2] J. Friedman, T. Hastie, and R. Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441, 2008.

[3] W. Leutzbach. *Introduction to the Theory of Traffic Flow.* Springer, 1988.

[4] T. Liebig, N. Piatkowski, C. Bokermann, and K. Morik. Predictive trip planning – smart routing in smart cities. In *Workshop Proceedings of the EDBT/ICDT 2014 Joint Conference*, 2014.

[5] T. Liebig, Z. Xu, M. May, and S. Wrobel. Pedestrian quantity estimation with trajectory patterns. In *Machine Learning and Knowledge Discovery in Databases*, volume 7524 of *Lecture Notes in Computer Science*, pages 629–643. Springer Berlin Heidelberg, 2012.

[6] L. Song, M. Kolar, and E. P. Xing. Time-varying dynamic bayesian networks. *Advances in Neural Information Processing Systems*, 22:1732–1740, 2009.

[7] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.