

Autonomous Accident Monitoring Using Cellular Network Data

Olof Görnerup

Swedish Institute of Computer Science
olofg@sics.se

Per Kreuger

Swedish Institute of Computer Science
piak@sics.se

Daniel Gillblad

Swedish Institute of Computer Science
dgi@sics.se

ABSTRACT

Mobile communication networks constitute large-scale sensor networks that generate huge amounts of data that can be refined into collective mobility patterns. In this paper we propose a method for using these patterns to autonomously monitor and detect accidents and other critical events. The approach is to identify a measure that is approximately time-invariant on short time-scales under regular conditions, estimate the short and long-term dynamics of this measure using Bayesian inference, and identify sudden shifts in mobility patterns by monitoring the divergence between the short and long-term estimates. By estimating long-term dynamics, the method is also able to adapt to long-term trends in data. As a proof-of-concept, we apply this approach in a vehicular traffic scenario, where we demonstrate that the method can detect traffic accidents and distinguish these from regular events, such as traffic congestions.

Keywords

Cellular networks, mobility, anomaly detection, emergency response, crisis management.

INTRODUCTION

Mobile communication networks consist of cells that each covers a geographic area. When a mobile phone is active it is linked to a specific cell - typically the one that provides the strongest signal - which occasionally changes as the mobile phone is moving. By combining statistics about the number of phones that are linked to each cell over time with the cells' geographic coverage, we can infer collective mobility patterns. We propose a method for capturing significant shifts in such patterns in order to monitor, detect and assess the impact of accidents and other critical events. The long-term aim is to develop robust, decentralised processes that run within the cellular network itself in order to provide such services. These services can in addition enable basic mechanisms for improving the robustness of cellular networks by enhancing their ability to adapt to rapidly changing and anomalous loads that may occur during crisis. Using cellular networks as massive sensor networks is advantages in several respects. Firstly, cellular networks are ubiquitous and cover most populated areas. Secondly, they are already in place and do not require major investments in infrastructure. Thirdly, they are commonly used, and therefore produce huge volumes of data. Leveraging this data we can maintain reliable statistical models and estimates of collective mobility patterns.

Our approach is based on a local measure that is approximately invariant on short time-scales under regular variations in mobility patterns. Long and short-term dynamics of this measure are estimated using Bayesian inference. This enables us to adapt to long-term trends as well as identify sudden changes as these cause the long and short-term dynamics to significantly diverge. The method is distributed as each network cell monitors local divergences. This, in combination with the large coverage of a cellular network, will give us a bird's eye view of the effect of an accident in terms of the geographic distribution of divergences. Such a view is valuable when allocating emergency resources, in particular if there are multiple simultaneous incidents, for instance due to a natural disaster. Having a distributed and dynamic view of anomalies may also prove invaluable when finding the most efficient way to redirect traffic or direct an evacuation, or when determining which routes are blocked and should be avoided by emergency services.

Mobility information in cellular network data has previously been used in several different areas, such as for classifying trajectories (Becker, Caceres, Hanson, Loh, Urbanek, Varshavsky and Volinsky, 2011), inferring

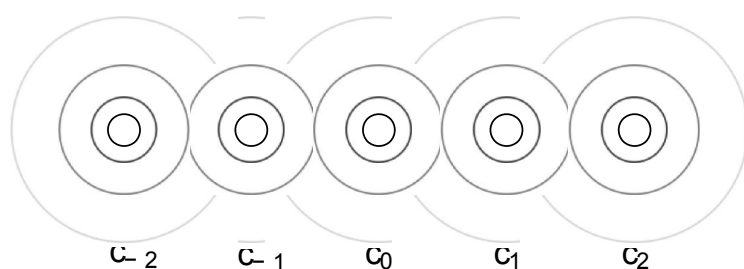


Figure 1. Four-lane highway covered by overlapping network cells c_i . Arrows indicate inflow of vehicles.

common routes (Görnerup, 2012), for location and trajectory prediction (Bhattacharya and Das, 1999; Yavas, Katsaros, Ulusoy and Manolopoulos, 2005), for determining modes of transportation (Sohn, Varshavsky, LaMarca, Chen, Choudhury, Smith, Consolvo, Hightower, Griswold and de Lara, 2006) or if a cell phone is stationary (Smith, Chen, Varshavsky, Sohn, and Tang, 2005), for mobility modeling and network tracking area optimisation (Chondronasiou and Papadia, 2011; Kreuger, Gillblad and Arvidsson, 2012), for modeling behaviour during emergencies (Chen, Zhai and Madey, 2011), and for studying wireless communication in vehicular traffic (Massey, W., and Whitt, W, 1994). The potential of utilising cell phone data in crisis management has also been demonstrated in (Bengtsson, Lu, Thorson, Garfield and von Schreeb, 2011), where cell usage and tower position data are successfully used to estimate population movements during a natural disaster and a cholera outbreak.

Cellular networks have also been subject to anomaly detection (Xie, Han and Tian, 2011), for instance in (Pawling, Chawla, and Madey, 2007), where clustering is used to detect anomalous data points as outliers in call data records, e.g. in terms of initiation time, duration and type of service. In contrast to their work, we consider anomalous states of statistical representations, which we believe is more robust due to a higher resilience to noise in data, and less sensitive to privacy intrusions. The approach by Pawling et al. is used in a general-purpose system for detecting emergencies and suggesting responses by using cell phone data in concurrence with agent-based modeling (Schoenharl and Madey, 2006). With respect to their work, our approach differs in that it is completely data-driven. Similarly, (Candia, González, Wang, Schoenhar, Madey and Barabási, 2008) consider collective mobility behaviour inferred from mobile communication data, and detect spatiotemporal anomalies using an approach based on percolation theory. Focusing on fraud detection, (Sun, Yu, Wu and Leung, 2004) instead propose an anomaly detection algorithm based on higher-order Markov models of mobility patterns. In comparison to their work, we instead propose a distributed and autonomous method that can be run within the cellular networks itself.

In this paper we will demonstrate the applicability of our approach by presenting a proof-of-concept within a specific vehicular traffic scenario. In this setting, an approximately time invariant measure is defined as local densities of cell phones linked to a network cell and its neighbouring cells. We show that this approach can be used to detect anomalies in the form of traffic accidents, and that these can be distinguished from regular variations in mobility patterns, e.g. due to traffic congestions. We will also show that the method can be used to quantitatively assess the impact of an accident in terms of affected subsequent traffic flow.

METHODS

Due to lack of empirical data - in particular involving traffic accidents - we test and evaluate our approach using simulations. The simulations are based on two components that model vehicular traffic dynamics and cellular network activity, respectively. The compound model is not intended to constitute a completely faithful representation of real traffic and network dynamics, but rather it enables us to provide a proof-of-concept for motivating more extensive studies and experiments. In order to decrease the risk of introducing model-dependent artifacts, we formulate and study a simple bare-bones scenario. We consider a stretch of a four-lane highway, where there are no slipways, exits or road crossings. At each end of the stretch, vehicles enter in both lanes with a given rate. The highway is evenly covered with network cells (further explained below), and the environment is completely homogenous in terms of geography, buildings etc. See Figure 1 for a schematic illustration of this scenario.

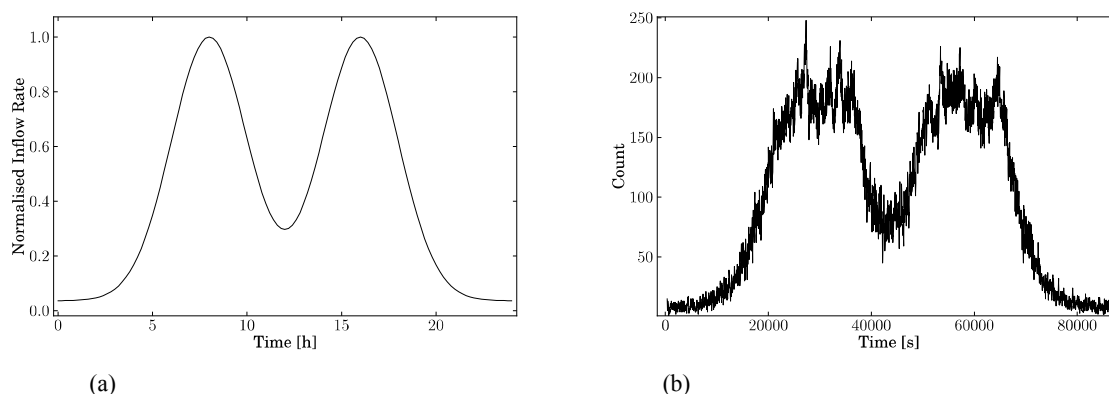


Figure 2. (a) Vehicle inflow modeled by a truncated mixture of two Gaussians. (b) Time series of the number of vehicles linked to a given cell, derived from simulation.

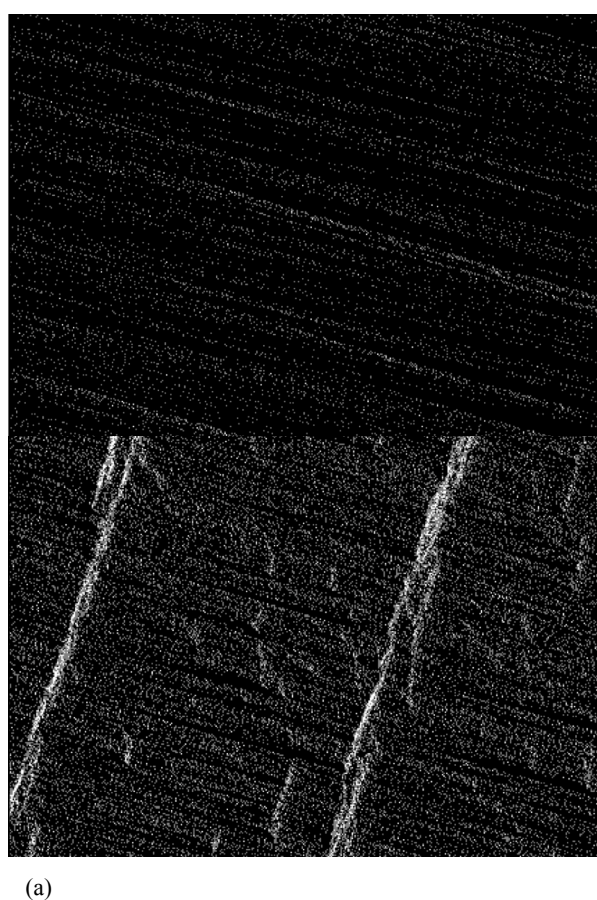


Figure 3. Space-time diagrams starting at (a) 17000 s. and (b) 31000 s. Traffic runs from left to right, where a site is white if both lanes are occupied by vehicles, gray if one lane is occupied, and black if both lanes are empty. Time runs from top to bottom. The diagrams show the highway traffic in one direction. Note in (b) the emerged traffic jams that form waves that travel in the opposite direction of traffic.

Vehicular traffic model

The vehicular traffic simulation is based on a space- and time discrete model introduced by (Nagel, Wolf, Wagner, and Simon, 1998). In the simulation each vehicle obeys a set of rules that are defined in terms of the state of the vehicle (lane occupied and velocity) and its local neighbourhood (distance to the closest vehicles and their velocities). The model manages to capture properties observed from empirical traffic measurements, where the spontaneous formation of traffic jams is the most relevant property for this study. As in real traffic, these

congestions emerge as waves that travel in the opposite direction of traffic.

Following the results in (Grabec, Kalcher and Svegl, 2008), we model the time dependent inflow of traffic by using a truncated mixture of two Gaussians, see Figure 2(a), in order to capture variations of inflow during 24 hours, such as low activity at night compared to high inflow rates at morning and evening rush hours. The purpose for modeling these variations is to determine how different hourly conditions affect our ability to distinguish accidents. A related objective of this study is to determine if it is possible to distinguish a common traffic jam from an accident using cellular network data. To further trigger traffic congestions at peak hours, we therefore extend Nagel et al.'s model by assigning each vehicle to one of two categories, where one category (that may be interpreted as buses and trucks) has a lower maximum velocity than the other (constituting regular cars). As exemplified in Figure 3, the model indeed exhibits spontaneous formation of traffic jams at peak hours. An accident is simply modeled by blocking one of the lanes in one direction at the centre of the highway stretch.

Cellular network model

In practice, network cells overlap due to local perturbations in signal strength, caused for instance by varying geographic topology or obstacles such as buildings and vegetation. We simulate these perturbations in signal strength by using a model introduced by (Erceg, Greenstein, Tjandra, Parkoff, Gupta, Kulic, Julius and Bianchi, 2006), which is based on experimental signal strength data that is fitted to a stochastic model. In the scenario we assume that cell phone usage is homogenous in all vehicles and that this usage is independent of the current vehicular traffic state. These are most likely false assumptions. One can for instance imagine that cell phones are used to a higher extent in traffic jams, or in the vicinity of major traffic accidents, than in free-flowing traffic. However, such variations are also likely to be advantageous for our purposes, as they result in signals that may be used to detect accidents more easily - signals that we consequently chose to discard in this study for sake of simplicity.

The cell used by a phone is known by the network under several circumstances, e. g. at connections and disconnections, when performing handovers between cells, and when changing location areas. Under the simplifying assumption that network events that provide information about vehicle cell usage are time invariant and vehicle-independent, we refrain from modeling possible network events in detail in order to avoid introducing a multitude of additional parameters and dependencies. Instead, a uniform sample of all vehicles generates events that signal their cell usage. In this way we approximate the cell information as seen by the network, modulo a scaling factor that reflects the probability that a phone generate network events, by counting the number of vehicles for which each cell provides the highest signal strength. The hourly variations of such counts for a cell are exemplified in Figure 2(b). This plot demonstrates the stochasticity that is due to varying vehicular traffic and signal strengths, as well as shows how traffic congestions are revealed as oscillations at peak hours.

Monitoring method

Our strategy is to use a statistical measure based on cell usage that is approximately invariant under hourly variations and traffic congestions, but that changes when an accident occurs. In the highway scenario, the following measure D fulfills this property: For each cell c_i , let D be the local distribution of vehicles linked to c_i and its neighbouring cells c_{i-1} and c_{i+1} . That is,

$$D = \{d_{i-1}, d_i, d_{i+1}\},$$

where d denotes relative vehicle density with respect to the three cells. This measure is approximately invariant since it constitutes a relative local load that is roughly constant in relation to raw vehicle counts as illustrated in Figure 2(b).

Bayesian estimators

We dynamically estimate D from data using recursive Bayesian inference (Kreuger, Gillblad and Arvidsson, 2012; Steinert and Gillblad, 2010), where prior estimates are influencing the current estimate. That is, at time t ,

$$d_j(t) = \frac{d_j(t - \Delta t)r + n_j(t)}{r + \sum_{k=i-1}^{i+1} n_k(t)},$$

where n_j denotes vehicle count associated with cell j and r is a parameter that determines the degree of influence of the prior distribution, i.e. the time-scale considered (r is also interpreted as the number of samples used to estimate the prior distribution). The initial distribution, $d_j(0)$, is set to uniform in order to achieve maximum

entropy. The estimate is updated each time a sufficient number of samples have been collected (this takes Δt seconds) in order to ensure reliable statistics. In this way we update D more frequently during rush hours than during night, for example.

Two measures are maintained in parallel, where one measure, D_L , captures long-term trends, and the other, D_S , captures short-term dynamics. Both measures are estimated as above, but differ in that D_L is less sensitive to recent dynamics. That is, $r_S \ll r_L$. Maintaining these two measures, we can capture sudden variations caused by accidents, by detecting differences in D_L and D_S , and at the same time adapt to long-term trends that are captured by D_L .

Divergence measure

The difference between the long- and short-term dynamics is quantified as the Kullback-Leibler (K-L) divergence, K , between D_L and D_S , where the former is considered to be the baseline (the "normal" state):

$$K(D_L; D_S) = \sum_i d_{L,i} \log \frac{d_{L,i}}{d_{S,i}}, 0 \log 0 \equiv 0.$$

K-L divergence is an information theoretic measure that is commonly used for comparing probability distributions. It is interpreted as the amount of information that is lost when approximating D_L with D_S . In this context K-L also constitutes a quantitative measure of the impact of an event such as an accident. When K reaches a given threshold value, an accident is considered to be detected.

RESULTS

In the following simulations we consider a 4 km long stretch of highway covered by 5 network cells, where the centres of the first and the last cells are located at the endpoints of the highway. 70% of the vehicles have a maximum velocity of 100 km/h, whereas 30% have a maximum velocity of 80 km/h. All vehicles are used when estimating D_L and D_S . Lowering the fraction of vehicles considered will qualitatively give the same results as those reported here.

Example

We begin by exemplifying how the measure estimates and their divergence typically evolve after an accident, as shown in Figure 4. The dynamics of the measures depend on the time of the day, where peak hours are characterised by larger oscillations of the short-term estimate due to waves of traffic congestions. This can be seen when comparing Figure 4(a) and (b), where the former displays less variation due to the lack of traffic jams. Despite the variations at peak hour, an accident causes significant divergence that can be detected, Figure 4(f), although the divergence is more distinct earlier in the day, as seen in Figure 4(e). Note that D_L is relatively stationary regardless of the time of day, Figure 4(c) and (d).

Evaluation

We systematically evaluate the method by performing multiple runs under different conditions. More specifically, we vary the start time in discrete steps, $t_0=10000m$ seconds for $m=1, 2, \dots, 8$, and the detection threshold, $s=0.01, 0.02, \dots, 0.05$ bits, and measure the following properties:

- The fraction of accidents that are detected.
- The fraction of false positives (when the detector triggers although no accident has occurred).
- The time it takes to detect an accident.

In practice, we perform each test as follows:

1. Set parameters t_0 and s .
2. Run the vehicular traffic simulation from t_0-2g to t_0-g seconds in order to avoid initialisation artifacts, where $g=30$ minutes.
3. Start updating the estimates and run the traffic simulation for another g seconds.
4. In half of the occasions, simulate an accident by blocking a lane in one direction.
5. Continue running the traffic simulation while updating D_L and D_S .
6. Abort the simulation if the divergence, K , reaches threshold s , or if the threshold is not reached within g seconds.

Using the above test, we can estimate the fraction of correct accident detections, the probability of false positives and the detection time. As indicated in the example given in Figure 4, the traffic conditions in terms of

congestions and lack thereof have an impact on these properties. Firstly, at very low vehicle inflow rates, $m=1, 7$ and 8 , accidents have no impact on the throughput of traffic, and therefore none of the accidents are detected. This is not due to that the method is not able to detect an anomaly, since there are no anomalies to detect. At the same time, there are no false positives since the threshold is above the noise floor at these start times. It is interesting to note the significant difference for $m=2$ and $m=7$, since the inflow at those times are roughly the same. The prior buildups of traffic are different though, as well as the higher order influx rates (a steep increase of rate, versus a steep decrease of rate) - both factors that affect the result. Secondly, *all* accidents are detected for $m=2$, whereas there are no false positives: an ideal situation due to a maximum vehicle inflow prior to the emergence of major traffic jams. The advantageous conditions are also reflected by the distinct increase of K in Figure 4(e). This start time also stands out when studying the estimated distributions of detection times for threshold $s=0.05$ bits, Figure 6, as it results in less variance than at other start times.

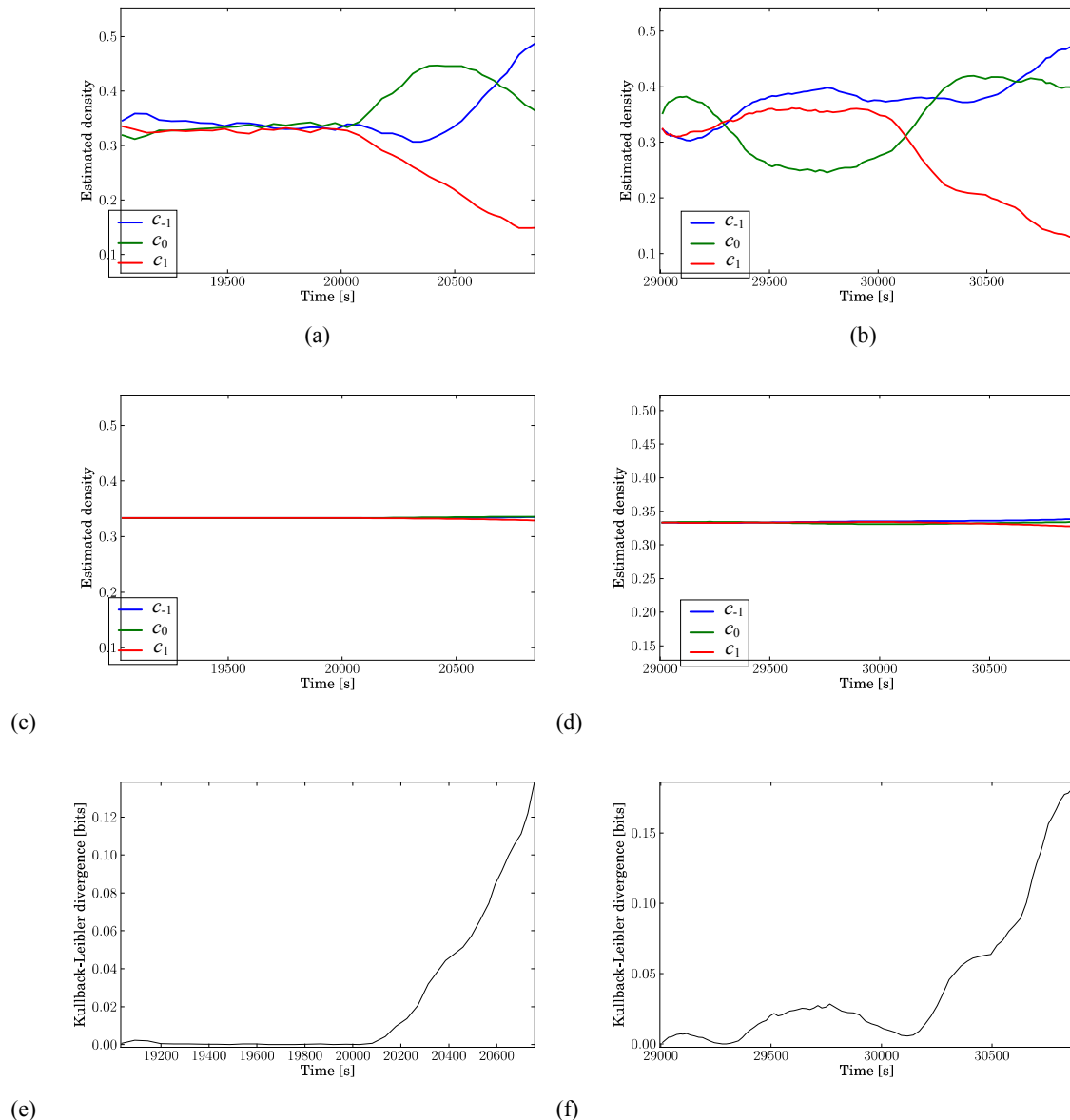


Figure 4. Evolution of short-term, (a) and (b), and long-term, (c) and (d), estimates, where three densities are shown. These densities are associated with the middle cell in Figure 1 and its two neighbouring cells. The Kullback-Leibler divergence between (a) and (c), and (b) and (d), are shown in (e) and (f), respectively. Accidents occurring at 20000 s (left) and 30000 s (right) cause the short and long-term estimates to diverge, as seen in (e) and (f).

Thirdly, the results at the remaining start times, for $m=3$ to 6, are more subtle as seen in Figure 5. All accidents are in these cases also detected, but as shown in Figure 5(b), the degree of false negatives is lower for $m=4$ due to fewer traffic congestions at midday. At the same time, the mean detection time is lower and on par with the detection time for $m=2$. In general, since detection times grow with the threshold, Figure 5(a), whereas the fraction of false positives decreases (with the exception of $m=2$), Figure 5(b), we face a tradeoff that is dictated by what is considered as reasonable expected detection times versus the risk of incorrectly detecting an accident. Again viewing the histograms in Figure 6, we see that the detection times can differ significantly for the same threshold. This is not surprising due to the strong stochastic components in the models, where large variations in dynamics have the consequence of smearing out the detection times.

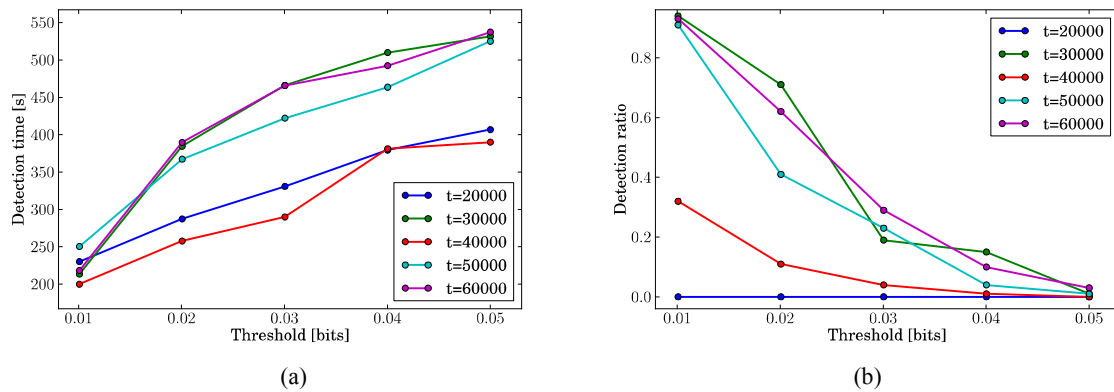


Figure 5. (a) Mean detection times and (b) fractions of false positives. 100 runs per data point.

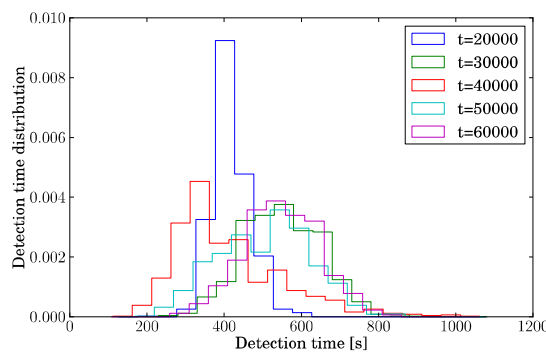


Figure 6. Histograms of detection times for threshold 0.05 bits. 1000 samples are used per histogram.

DISCUSSION

Even though we consider a bare-bones scenario that we model using several simplifying assumptions, the results reported here are promising and point to the possibility of employing data from cellular networks for detecting anomalies caused by critical events such as vehicular traffic accidents. As already pointed out, the study is intended to serve as a proof-of-concept, and the results should be interpreted with this in mind, as the quantitative outcomes of our evaluations will most likely not directly translate to a real scenario. Taking this into consideration, we would like to underline that the specific minutes of response times reported here, for instance, are not the main take-home-message. Results of relevance are instead that accidents *are* detectable, and that the detection time is on the order of minutes rather than hours.

The approach would improve crisis response and management by enabling existing cellular networks to serve as ubiquitous sensor networks that require little additional investments in infrastructure and deployment. As the approach is distributed, autonomous and to a certain degree redundant, it is also robust in the sense that portions of the network may fail without disrupting the monitoring system as a whole. Although the granularity is relatively coarse (given by the geographic area covered by a cell) compared to GPS data, for instance, the method leverages on the huge amounts of data produced in the network in order to infer large-scale mobility

patterns. Furthermore - again in comparison to methods based on satellite data (Voigt, Kemper, Riedlinger, Kiefl, Scholte and Mehl, 2007) - the proposed approach is also valid underground, such as in a subway system, and less sensitive to weather conditions (especially compared to methods based on image processing). For example, imagine a road tunnel, perhaps several kilometers long, that is supported by a cellular network (this is a scenario very similar to that depicted in Figure 1). By employing the proposed approach we may autonomously monitor the tunnel without having to install additional sensors or communication equipment, but by utilising the sensors that are already present: the cellular network.

So far we have solely treated a specific scenario - a stretch of four-lane highway - and a specific application - monitoring and detecting traffic accidents. However, our approach is very general; identify a measure that is approximately constant on short time-scales under regular circumstances, estimate its dynamics on short as well as long long time-scales, and monitor the divergence between these estimates. Clearly, such divergences can be due to any types of anomalies causing unusual dynamics. The method is in other words blind to the underlying cause of the divergence, and may very well be applicable in other scenarios, in and beyond crisis response and management. Another aspect to which the method is blind is that of time-scale. Here the short- and long-term estimators capture dynamics on the order of minutes and hours, respectively, but it could as well be days, weeks or months etc. A relevant question could for instance be; is the traffic flow this month unusual compared to the traffic flow the same month a year ago? Being able to consider several time-scales in this way opens up yet other potential applications areas, where it is relevant to monitoring and detecting subtle and nonobvious changes that happen over months or even years.

When using cellular network data, privacy must be taken into consideration as location data may be perceived as sensitive. Since our approach is based in statistical models - aggregated information in the form of local densities - references to individuals or small groups of people and their whereabouts are not necessary, nor desirable from a statistical point of view. In fact, the accuracy and reliability of our approach are intertwined with anonymisation, in the sense that a large number of samples are a prerequisite for the method to work properly.

CONCLUSIONS

We have proposed an autonomous method for monitoring and detecting significant shifts in mobility patterns reflected in cellular network data. The method is based on detecting divergences between short and long-term dynamics of a measure that under regular circumstances is approximately time-invariant on short time-scales. We demonstrate this approach by studying a vehicular traffic scenario constituted by a stretch of four-lane highway. Using a compound model constituted by a vehicle traffic model and a cellular network model, we show that the method is capable of detecting accidents and distinguish these from regular traffic congestions, where the detection time and the likelihood of false alarms are found to be dependent on hourly variations in terms of inflow rate and presence of congestions. This proof-of-concept motivates further studies in more extensive scenarios, including empirical tests in a real-world setting.

ACKNOWLEDGMENTS

This work was funded by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

REFERENCES

1. Becker, R. A., Caceres, R., Hanson, K., Loh, J. M., Urbanek, S., Varshavsky, A. and Volinsky, C. (2011), Route classification using cellular handoff patterns, *Proceedings of the 13th international conference on Ubiquitous computing*, 123-132.
2. Bengtsson, L., Lu, X., Thorson, A., Garfield, R. and von Schreeb, J. (2011) Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data : A Post-Earthquake Geospatial Study in Haiti, *PLoS Medicine*, 8, 8, e1001083.
3. Bhattacharya, A. and Das, S. (1999) LeZi-update: An Information-Theoretic Approach to Track Mobile Users in PCS Networks, *MobiCom 99*, 1-12.
4. Candia, J., González, M. C., Wang, P., Schoenhar, T., Madey, G. and Barabási, A.-L. (2008) Uncovering individual and collective human dynamics from mobile phone records, *J. Phys. A: Math. Theor.*, 41, 224015.
5. Chen, F., Zhai, Z. and Madey, G. (2011) Dynamic adaptive disaster simulation: developing a predictive model of emergency behavior using cell phone and GIS data, *Proceedings of the 2011 Workshop on Agent-*

Directed Simulation, ADS '11, 8, 5-12.

6. Chondronasiou, M. and Papadia, P. (2011) Mobility Modeling In Cellular Networks, Masters Thesis, KTH, Royal Institute of Technology.
7. Erceg, V., Greenstein, L. J., Tjandra, S. Y., Parkoff, S. R., Gupta, A., Kulic, B., Julius, A. A. and Bianchi, R. (2006) An empirically based path loss model for wireless channels in suburban environments, *IEEE J.Sel. A. Commun.* 17, 7, 1205-1211.
8. Grabec, I., Kalcher, K. and Švegl, F. (2008) Statistical modelling of traffic flow rate. *SIRWEC 2008*, Prague, 14-16 May 2008.
9. Görnerup, O. (2012) Scalable Mining of Common Routes in Mobile Communication Network Traffic Data, *Proceedings of the 10th International Conference on Pervasive Computing*, 7319, 99-106.
10. Kreuger, P., Gillblad, D., and Arvidsson, Å. (2012) Zero configuration adaptive paging (zCap), *IEEE 76th Veh. Technol. Conf.*, 3-6 Sept 2012, Québec, Canada.
11. Massey, W., Whitt, W. (1994) A stochastic model to capture space and time dynamics in wireless communication systems. *Probability in the Engineering and Informational Sciences* 8, 541–569.
12. Nagel, K., Wolf, D. E., Wagner, P., and Simon P. (1998) Two-lane traffic rules for cellular automata: A systematic approach, *Phys.Rev.*, E58, 1425–1437.
13. Pawling, A. Chawla, N. V. and Madey, G. (2007) Anomaly detection in a mobile communication network. *Comput. Math. Organ. Theory* 13, 4, 407-422.
14. Smith, I., Chen, M., Varshavsky, A., Sohn, T., and Tang, K.P. (2005). Algorithms for Detecting Motion of a GSM Mobile Phone. *Workshop on Location Awareness and Community, at ECSCW 2005*.
15. Sohn, T., Varshavsky, A., LaMarca, A., Chen, M.Y., Choudhury, T., Smith, I., Consolvo, S., Hightower, J., Griswold, W.G., de Lara, E. (2006) Mobility Detection Using Everyday GSM Traces, *UbiComp 2006*, 4206, 212–224.
16. Steinert, R. and Gillblad, D. (2010) Long-Term Adaptation and Distributed Detection of Local Network Changes, *IEEE Global Telecomm. Conf. GLOBECOM*, 1-5.
17. Sun, B., Yu, F., Wu, K. and Leung, V. (2004) Mobility-based anomaly detection in cellular mobile networks. *Proceedings of the 3rd ACM workshop on Wireless security*, 61-69.
18. Schoenharl, T. and Madey, G. (2006) WIPER: Leveraging the cell phone network for emergency response, *International Journal of Intelligent Control and Systems*, 11.
19. Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., and Mehl, H. (2007), Satellite Image Analysis for Disaster and Crisis-Management Support, *IEEE Transactions on Geoscience and Remote Sensing*, 45(6-1), 1520-1528.
20. Xie, M., Han, S., Tian, B. and Parvin, S. (2011). Anomaly detection in wireless sensor networks: A survey, *J. Netw. Comput. Appl.* 34, 4, 1302-1325.
21. Yavas, G., Katsaros, D., Ulusoy, and Manolopoulos, Y. (2005) A data mining approach for location prediction in mobile environments. *Data Knowl. Eng.* 54, 121–146.