Application of Tracking Signals to Detect Time Series Pattern Changes in Crime Mapping Systems

By

Wilpen L. Gorr
H. John Heinz III School of Public Policy and Management
Carnegie Mellon University

And

Shannon A. McKay
Department of Urban Studies and Planning
Massachusetts Institute of Technology

Abstract

Tracking signals are widely used in industry to monitor inventory and sales demand. These signals automatically and quickly detect departures in product demand, such as step jumps and outliers, from "business-as-usual". This paper explores the application of tracking signals for use in crime mapping to automatically identify areas that are experiencing changes in crime patterns and thus may need police intervention.. Detecting such changes through visual examination of time series plots, while effective, creates too large a work load for crime analysts, easily on the order of 1,000 time series per month for medium-sized cities. We demonstrate the so-called smoothed-error-term tracking signal and carry out an exploratory validation on 10 grid cells for Pittsburgh, Pennsylvania. Underlying the tracking signal is an extrapolative forecast that serves as the counterfactual basis of comparison. The approach to validation is based on the assumption that we wish tracking signal behavior to match decisions made by crime analysts on identifying crime pattern changes. We present tracking signals in the context of crime early warning systems that provide wide area scanning for crime pattern changes and detailed drill-down maps for crime analysis. Based on preliminary results, the tracking signal is a promising tool for crime analysts.

INTRODUCTION

Police know the current crime patterns in their jurisdictions and accordingly allocate manpower to precincts and shifts, target patrols to hot spots, and take other tactical actions. What is less well known to police is how crime patterns are changing, so that police can reallocate manpower in response to changes. We learned this lesson in the early 1990s when we built a crime mapping system for the Pittsburgh, Pennsylvania Bureau of Police under a Drug Market Analysis Program (DMAP) grant funded by the National Institute of Justice. Many times our DMAP crime mapping system detected enforcement-induced displacement of street-level drug dealing before narcotics detectives were able to do so. Follow-up surveillance of new drug dealing locations detected by our system always proved the maps to be right.

From this experience we learned the value of building crime early warning system (CEWS) maps.   These maps display crime changes to provide a jurisdiction-wide scan for areas needing changes in tactical deployment of police. Used on an interactive basis in a geographic information system, the maps provide drill-down to areas of high change to provide detailed, diagnostic information. We provide example maps below, but before proceeding to them, it is important to distinguish two types of change: experienced and forecasted change.

Experienced change is the sort mentioned above, which has the objective of quickly detecting any sort of crime innovation (departures from business-as-usual crime patterns), such as crime displacement in response to enforcement.

Underlying analytic problems are 1) to provide counterfactual (business as usual) forecasts as the basis of comparison for the most recent historical crime data and 2) to sort out true pattern changes from random variations. More is on these issues below. Detecting experienced change is a major activity in Comstat meetings (see Henry and Bratton 2002)[1].

The second type of change - not the subject of this chapter - is forecasted change, which provides some capacity for crime prevention. Recently there has been success on developing crime forecasting as an applied research field (e.g., see Gorr and Harries, 2003 which introduces a special section on crime forecasting in the *International Journal of Forecasting*). Extrapolation of crime seasonality and time trend one month ahead have proven to be accurate enough for use in tactical deployment given adequately high crime rates in areas investigated (Gorr, Olligschlaeger, and Thomson 2003).

Our purpose in this paper is to introduce and examine tracking signals as a potential tool for crime analysts for automatically detecting crime innovations. We undertake an exploratory empirical validation of one of the best tracking signals. We have not seen any validation studies in the literature using real data such as used here. All have relied on simulated data with known pattern changes for validation. Instead, we use judges (ourselves) to visually identify pattern changes. The next section provides examples of CEWS maps to provide the context for tracking signals (and crime forecasting) as tools for use by crime

---

[1] Note that this paper pursues experienced change relative to geographic areas, such as grid cells or census tracts. Another important pattern to establish, as an innovation, may cut across several geographic areas and is the identification of a serial criminal. In this case, analysis surrounds the matching of physical descriptions and modus operandi of perpetrators.

analysts in mapping crimes. The third section makes the case that an automated approach is needed to detect experienced crime pattern changes. The fourth section of this chapter briefly reviews tracking signals. Following that is a section on our research design for validation, followed by a section on results, and then a conclusion.

Crime Early Warning System Maps

Next, we provide examples of CEWS maps. Such maps appear similar in format whether using experienced or forecasted change. Thus while we do not have good example maps for experienced change at this time, the ones provided for forecasted change next are representative of change maps in general.

Figure 1 is a CEWS map for Pittsburgh, Pennsylvania displaying one-month-ahead crime forecasts where the areas are uniform grid cells 4,000 feet on a side (Gorr, Olligschlaeger, and Thomson 2003). The plotted values are forecasted changes in part 1 property crimes in December made at the end of November in a particular year. Increasingly dark solid-fill shading shows areas of increasingly larger forecasted increases and increasingly dark cross-hatching shows areas of increasingly larger forecasted decreases. While there are 103 grid cells, only nine have forecasts of sizable increases and of those only two have large increases (grid cells 61 and 77). Thus crime analysts would likely start working with the two worst cases, and then proceed to the other seven.

CEWS includes drill-down to individual crime points of the most recent month – either for the crime type of the grid cells (part 1 property crimes) or

corresponding leading indicator crimes (such as criminal mischief and disorderly conduct). Figure 2 is a drill-down (zoom in) to grid cell 77 showing crime points for two part 1 property crime types, burglary and larceny, in November. Clearly, there are hot spot clusters for both crime types. Based on an assumption of persistence for the hot spots (e.g., Block 1995, Harries 1999, Liu and Brown 2003), and a study of corresponding crime reports and modus operandi data (e.g., place of entry, time of day, etc.), crime analysts can suggest places and times to patrol hot spot areas within grid cells.

NEED FOR AUTOMATED DETECTION

A problem with attempting to identify crime time series pattern changes for current conditions is that the analyst must examine time series plots of about five years length each month. Analysts have to account for regular noise versus departures from established time trend patterns such as a sudden discrete change (step up or down) or a turning point (e.g., change from a decreasing time trend to an increasing trend). This work can be done by visual examination, but generates an unacceptably-large workload because analysts must work with relatively small geographic areas, such as grid cells or census tracts. For

Figure1: Early Warning System with Forecasted Change in Serious Property Crimes For December Made at the End of November.
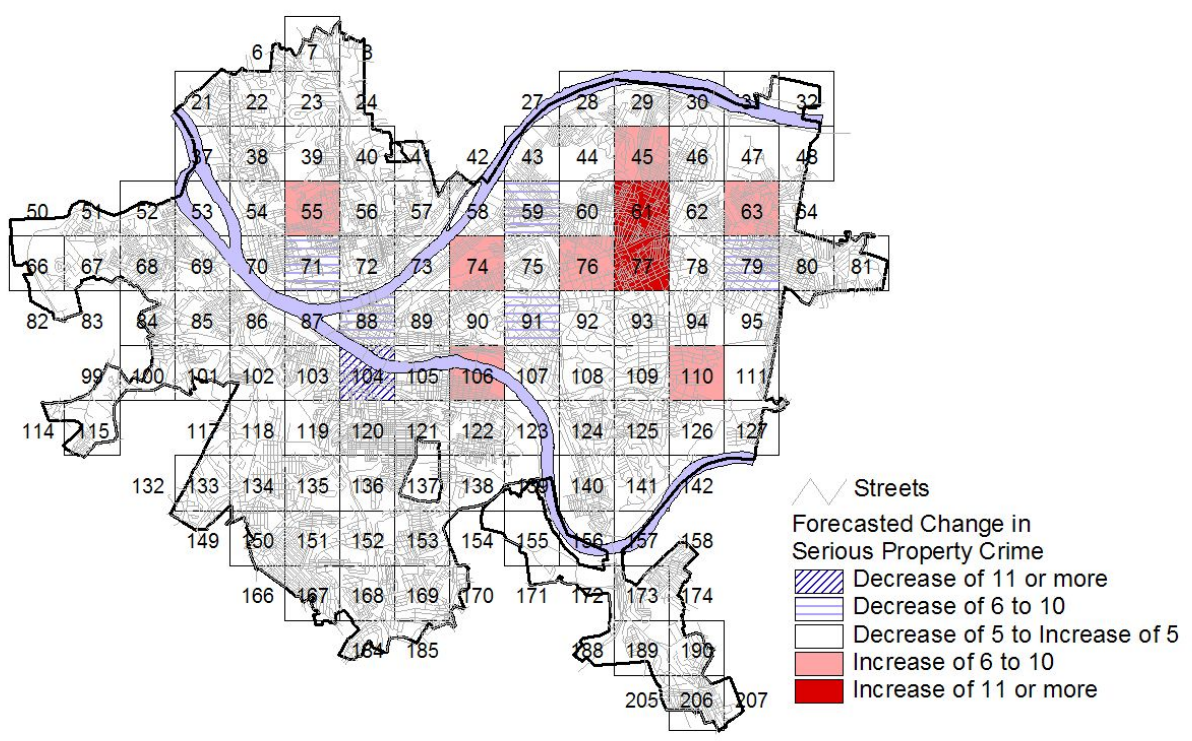

Figure 2: Zoom-In to Grid Cell 77 to View November Crime Points.

Buildings and Streets map layers provided by Pittsburgh City Planning Department

example, in Pittsburgh, there are approximately 100 grid cell areas to examine

and at least 10 crimes of interest, yielding roughly 1,000 crime series plots to

generate and examine each month. Clearly it is infeasible to implement pattern

change detection with visual examination. This is where tracking signals come

into play. They automatically flag exceptional time series.

Time series tracking signals are widely used by businesses for sales

forecasting and inventory control to generate exception reports of time series that

have likely deviated from their historical time trends. Next is a brief review of

tracking signals.

TRACKING SIGNALS

An approach to evaluating a phenomenon at a point in time is to make a

counterfactual forecast for the point, which predicts the point under business-as-

usual conditions. Then a tracking signal can be established, based on the actual

crime data point and in reference to the corresponding counterfactual, so that if

the tracking signal exceeds a selected control limit, an exception report is tripped

for a potential time series pattern change. We use extrapolative time series

forecasts to make counterfactual forecasts; namely, the most accurate

extrapolative forecast method as determined by Gorr, Olligschlaeger, and

Thomson (2003) for one-month-ahead crime forecasts. This is Holt exponential

smoothing with smoothing parameters optimized (see Bowerman and O'Connell

1993, pp. 400-403) and using time series data deseasonalized with multiplicative

seasonal factors estimated from jurisdiction-wide data and by classical decomposition (see Bowerman and O'Connell 1993, pp. 355-368). Thus, business-as-usual is defined to be a time series pattern following a smoothed linear time trend (straight line fitted to the time trend, placing most weight on the most recent data points) and monthly seasonal factors such as 1.25 (25 percent higher seasonal effect) or 0.80 (20 percent lower seasonal effect). The counterfactual forecast extends the estimated time trend ahead to the point being analyzed and applies the corresponding seasonal multiplier, using all prior data to estimate trend and seasonality.

Tracking signals generally are ratios in which the numerator is a sum or weighted sum of counterfactual forecast errors that has an expected value of zero when time series patterns (time trend and seasonality) are stable. When there is a pattern change, such as a step jump or turning point, the numerator moves away from zero. The denominator's purpose is to normalize by the long-term average variability of forecast errors. Of the common tracking signals, the smoothed error signal due to Trigg (1964) is a good choice for practitioners (McClain, 1988).   The equations are as follows:

$$E_t = \alpha_1 e_t + (1- \alpha_1)E_{t-1} \tag{1}$$

$$MAD_t = \alpha_2|e_t| + (1- \alpha_2)MAD_{t-1} \tag{2}$$

$$T_t = |E_t/MAD_t| \tag{3}$$

where

MAD = mean absolute deviation of forecast errors

E_t = smoothed forecast error

T = tracking signal

t = month being evaluated

e_t = counterfactual forecast error

α_1 = smoothing factor for numerator

α_2 = smoothing factor for denominator

We implement this signal with smoothing parameter values as suggested by McClain: α_1=0.40 for the smoothed sum of errors for the numerator (in order to quickly detect pattern changes) and α_2=0.05 for the denominator of smoothed mean absolute deviations of forecast errors. The initial value for E_0 is assumed to be 0, so that there is a burn-in period during which the tracking signal adapts to the actual pattern and forgets the initial value. In addition to computing the tracking signal, the analyst must also choose critical values which, if exceeded, trip an exception report. We make the critical value an experimental treatment, trying a range of critical values in an attempt to tune tracking signal behavior to match crime analysts' judgment on crime pattern changes.

These equations are easily implemented in a spreadsheet package for experimentation, but normally would be programmed to work automatically within a CEWS. Figure 3 is an example of equations 1-3 applied to monthly time series data for 911 drug calls in grid cell 120 of Figure 1. Each tracking signal value

has five years of historical data behind it in order to estimate corresponding

counterfactual forecast models.

Marked for comparison purposes are two pattern changes and an outlier.

The actual and forecasted crime levels have been rescaled to match the vertical

scale of the tracking signal. When the tracking signal crosses above the control

limit line, it issues (trips) an exception report, warranting analysis of this time

series. As J. McClain (1988, p. 563) states "A perfect tracking signal

Figure 3: Sample Tracking Signal for 911 Drug Calls in
Grid Cell 120 with Marked Pattern Changes and Outlier.
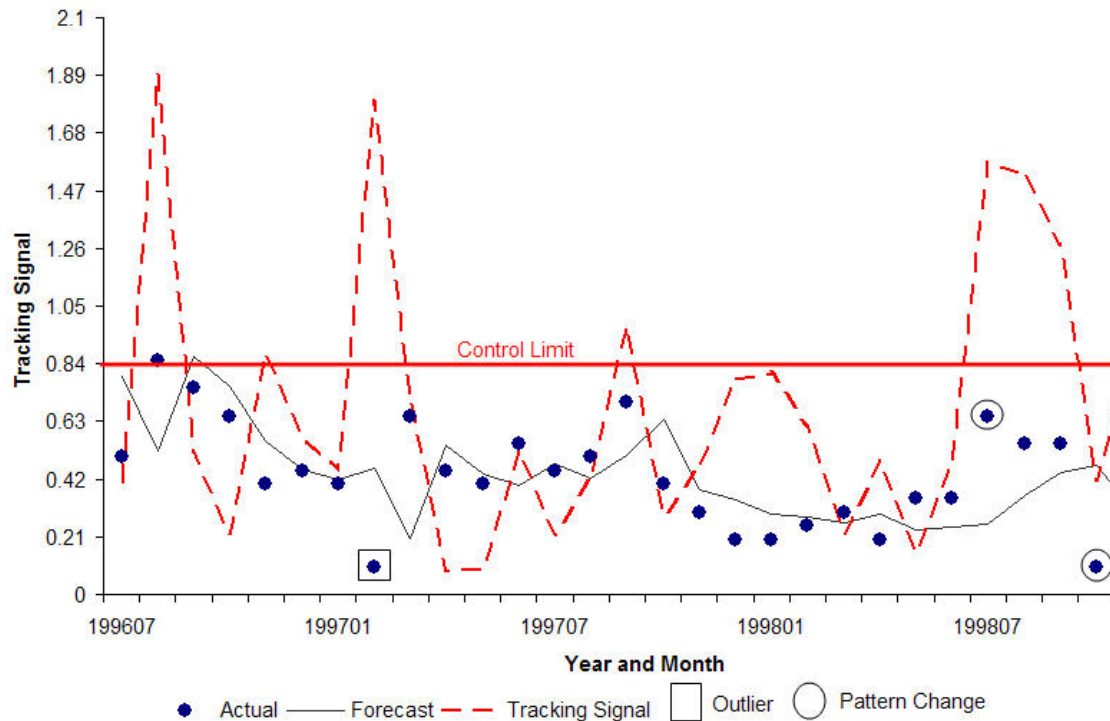
would detect an out-of-control forecast (i.e., a time series pattern change)

immediately, and would never give a false alarm." Of course, this is not possible,

so in Figure 3 the reader can see false positives (the first and third trips), but also

actual positives detected immediately (the second and fourth trips), and a delay in detecting an actual positive (the last data point which appears as if it would be detected if one more data point were available).

RESEARCH DESIGN

This section addresses the question of whether tracking signals really perform well for detecting changes in crime series patterns. We must account for false positives and determine if tracking signals reduce workloads adequately. We have not seen any attempts in the literature to validate tracking signals with actual data, as in Figure 3. All validations appear to have used simulated data with known pattern changes and outliers. It is very desirable, however, to use actual data in order to assess value in a given context; namely, will tracking signals adequately reduce workload and not miss actual positives? Thus, we assumed that the purpose of tracking signals is to match behavior of trained, human judges (crime analysts), and simply automate their decisions on pattern changes and outliers.

We did not have the resources to embark on a full-scale validation; hence, we decided to carry out an exploratory study to determine the feasibility of our approach and provide preliminary results. We chose 10 crime time series from the Pittsburgh grid system of Figure 1. They consist of a variety of crime types with five time series having pattern changes and the other five not having any. It is important to include time series with no pattern changes to assess false positive rates.

Both authors independently marked-up each of the time series for pattern changes and outliers, as in Figure 3, under the guideline that we would only mark those that are large and obvious. We then compared results and reconciled differences. One of us had merely admitted some smaller pattern changes in interpreting "large and obvious". The result was 18 instances of pattern changes or outliers in five of the time series used in our analysis.

Our treatment of the smoothed signal tracking signal was to use it with a variety of control limits, searching for the control limit that best matches detection of our judged, true pattern changes. After some trial and error, we decided to use values of 0.84, 1.05, 1.26, and 1.47. This range starts at a low value (0.84) that detects most of the actual positives, but at the cost of tripping many false positives (false alarms). At the other extreme (1.47), there are fewer detections of actual positives, but also many fewer false positives.

RESULTS

We applied equations 1-3 on the 10 time series over the 36 month period in which counterfactual forecasts were made. In reporting results, we decided to exclude the first six months of tracking signals for burn-in so that the tracking signal could forget arbitrary initial values and start tracking correctly. Hence there were 10 time series times 30 months each for a total of 300 signal values estimated. Also, this translates to 300 time series plots that a crime analyst would have had to examine to accomplish the same task.

We define an exception report "epoch" to be the total number of time periods that the tracking signal is above its control limit, including the first month that it trips. We assume that the crime analysis protocol is that the crime analyst must investigate each time series plot and corresponding crime maps for each month of epochs. Hence the count of all epoch months is a measure of the work load that the crime analyst would have to do when using tracking signals. The comparison without a tracking signal is 300 or 10 per month.

Table 1 is the result of our research. For a control limit of 0.84, the tracking signal detects 17 (94%) of the 18 actual positives, which appears to be quite good. It also does so with no lag or one period lag. The cost is, that of the average total of 4 time series per month to be examined (instead of 10), 2.9 are false positives. At the other extreme, with a control limit of 1.47, only 11 (61%) of the actual positives are detected, but the total workload per month is down to 1.6 time series, 1 of which is a false positive. The number of false positives falls quickly between the first two control limits in Table 1 and then flattens out.

Table 1.
Final Results on Validation Research

Control Limit	True Positives Detected	Average Workload (Time Series/Month)	Average False Positives (Time Series/Month)
0.84	17 (94%)	4.0	2.9
1.05	13 (72%)	2.8	1.9
1.26	12 (67%)	2.1	1.4
1.47	11 (61%)	1.6	1.0

We believe that these results are promising because they show a 60% work reduction for the most stringent case and up to an 84% work load reduction for the least stringent case.

CONCLUSION

This paper has discussed crime early warning systems and introduced an application of tracking signals for detecting experienced time series pattern changes in crime maps. The basis of the tracking signal is information obtained from counter-factual forecasts for each point examined. These are forecasts providing business-as-usual estimates for a point in time, as if no pattern changes existed. The tracking signal automates detection of pattern changes by matching the decisions of crime analysts as to what data points constitute the start of a new time series pattern. We varied the control limit of the tracking signal, making it more or less sensitive to information in the time series data in attempting to tune the tracking signal to match crime analysts' decisions.

In future work it will be necessary to take a large sample of time series, have crime analysts mark them up for pattern change points and outliers, and rerun the research study. Additional tracking signals may be tried, as well as varying the tracking signal numerator's smoothing factor (which we did not do) for further tuning and attempting to improve performance..

ACKNOWLEDGEMENTS

REFERENCES

Block, C. (1995). STAC hot spot areas: A statistical tool for law enforcement decisions in Block, C., Dabdoub, M, Fregly, S. [eds.] Crime analysis through computer mapping, Washington DC: Police Executive Research Forum.

Bowerman, B.L., O'Connell, R.T. (1993). *Forecasting and time series: an applied approach*. Belmont, CA, Duxbury Press.

Gorr, W.L., Harries, R. (2003). Introduction to crime forecasting. *The International Journal of Forecasting, Special Section on Crime Forecasting 16,551-555.*

Gorr, W.L., Olligschlaeger, A.M., Thompson, Y. (2003). Short-term time series forecasting of crime. *The International Journal of Forecasting, Special Section on Crime Forecasting 16, 579-594.*

Harries, K. (1999). Mapping crime: principle and practice, Washington DC, U.S. Department of Justice, National Institute of Justice Research Report, 112-118.

Henry, V.E., Bratton, W.J. (2002), The Compstat paradigm: management accountability in policing, business and the public sector. Flushing, NY: Looseleaf Law Publications.

Liu, H., Brown, D.E. (2003). Criminal incident prediction using a point-pattern-based density model. *The International Journal of Forecasting, Special Section on Crime Forecasting 16, 603-622.*

McClain, J.O. (1988). Dominant Tracking Signals. *The International Journal of Forecasting* 4, 563-572.

Trigg, D. W. (1964). Monitoring a forecasting system. *Operational Research Quarterly* 15, 271-274.

BIOGRAPHIES

Wilpen L. Gorr is Professor of Public Policy and Management Information Systems at the H. John Heinz III School of Public Policy and Management, Carnegie Mellon University and past editor of the *International Journal of Forecasting*. His research on forecasting, geographic information systems, and crime analysis has been supported by a series of grants from the U.S. National Institute of Justice and the Sloan Foundation. His publications have appeared in *International Journal of Forecasting*, *Management Science*, *Geographical Analysis*, *Journal of Geographical Systems, Criminology, Journal of Criminal*

Justice, and other journals.

Shannon A. McKay is a doctoral student in the Department of Urban Studies and Planning at Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts. Her research interests include urban sprawl, land use planning, growth management issues, education finance, as well as, fostering city and suburban cooperation. Shannon has a Master's of Science in Public Policy and Management from the H. John Heinz III School of Public Policy and Management at Carnegie Mellon University in Pittsburgh, Pennsylvania. Prior to attending Carnegie Mellon, she worked for 3 years at the Urban Institute in Washington, DC on education policy issues. She has interned at the U.S. General Accounting Office (GAO), as well as, the Urban Land Institute in Washington, DC. Her undergraduate degree is in economics from Vassar College.