December 2018

Coding Suspicion

Arthur Crivella[a], Wesley M. Oliver[b,1] and Morgan Gray[c]

[a]*Partner, Crivella Technologies, Ltd., Pittsburgh, PA*

[b]*Professor of Law, Duquesne University*

[c]*Third Year Law Student, Duquesne University*

Abstract. Legal standards for suspicion involve seemingly limitless possible factors, leaving them vague and subject to concerns of illegitimate biases by decision makers. Beginning with the relatively small number of factors present in drug interdiction stops, a model can be developed that not only predicts judicial behavior but the odds of discovering drugs. This technology will require legislatures or judges to begin the process of determining what numerical threshold of suspicion justifies investigatory detentions and searches.

Keywords. probable cause, reasonable suspicion, end user vernacular, vector regression

Crivella Technologies, in conjunction with researchers at Duquesne University School of Law, is developing a prototype for assessing whether officers performing drug interdiction stops have adequate suspicion for a search or prolonged detention. This technology holds the potential to improve the accuracy – and decrease implicit biases – in the on-the-scene decisions officers must make daily.

1. Unpredictable and Unreliable Legal Standards for Assessing Suspicion

Assessing suspicion is essential to law enforcement. American law requires police officers to have probable cause to conduct a search for evidence or arrest a suspect. With a lower quantum of suspicion – something the law calls reasonable suspicion – an officer may briefly detain an individual, and the car he is driving, for certain types of investigations. An officer, with this lesser level of concern, may question the detained suspect or have a drug dog sniff a motorist's car. [1]

The seemingly limitless number of factors upon which an officer may rely – and the low threshold requirements for permitting a search – provide little objective guidance for officers, or judges reviewing their decisions to search or arrest. Oliver Wendell Holmes, Jr., famously said that the law is a prediction of what a judge will do. [2] Multi-factored legal standards like probable cause or reasonable suspicion lack meaningful predictability, except when applied to circumstances identical to those judges have previously ruled upon.

Further, and perhaps more importantly, the conclusions of judges that particular sets of facts are sufficiently suspicious to justify an investigatory detention or search are never supported by empirical evidence. While police officers are forbidden to use "hunches," in deciding when to search, the limits of human ability prevent judicial findings of probable cause from being anything more than hunches with footnotes.

[1] Corresponding Author: Wesley M. Oliver, Duquesne Law School, 900 Locust Street, Pittsburgh, PA 15282, E-mail: oliverw@duq.edu.

Advanced cognitive computing, however, can be used to determine whether judges are attributing the appropriate weight to any particular suspicious factor, or combination of suspicious factors. Modern technology further permits a mean of applying our standards for assessing suspicion in a way that minimizes the implicit racial bias human actors are unaware they possess.

2. Reasonable Suspicion in Drug Interdiction Stops as an Ideal Model for Machine Learning of Legal Tests

The development of technology that aids in the assessment of suspicion must begin with a manageable subcategory of cases, one that contains a discrete number of possible suspicious factors. Drug interdiction stops provide such a starting point.

A large number of these cases exist. Officers must decide daily whether they will search motorists stopped for an ordinary traffic offense. Specialty units within police departments essentially do nothing else.

A necessarily small number of factors could give an officer a basis for further investigation in this context. During the brief time it takes an officer to write a traffic citation, and wait for a dispatcher to inform him whether the motorist has any outstanding warrants, he or she must determine what, if any, level of suspicion exists. Short of actually seeing or smelling illegal drugs themselves, the short list of bases of suspicion includes: observing a driver's nervousness, smelling air fresheners or talcum powder designed to disguise other smells, learning that the motorist reports he is traveling to a location different than one reported by his passenger or identified on a car rental agreement.

These factors are not binary and a meaningful model for assessing reasonable suspicion must account for this complexity. Each will exist in degrees and be described using very different terms. A smell can be strong or faint. Nervousness could be identified by sweating, stuttering, or a much more vague claim of anxiousness – and these symptoms could be explained by weather, physical impediments, or the manner of the officer in dealing with the motorists. A driver traveling north of Manhattan reporting that he is traveling to Maine offers a story somewhat inconsistent with his passenger who claims they are traveling to Vermont, but quite inconsistent with a rental agreement identifying Miami as the destination, and potentially consistent with a claimed destination of Boston.

There is, nevertheless, a limited universe of the circumstances that could justify suspicion of drug trafficking in an ordinary traffic stop. There are cases in which officers have reported rather unique circumstances that they claim raised their concern – playing loud gospel music for instance – but most efforts to justify either a search, or continued detention awaiting a drug dog, fit into a fairly small group of categories.

Modern computers have substantial advantages over humans in applying legal standards of suspicion. Computers obviously have faster reading capacity and better memories than any person. Machines can provide answers more consistent with existing case law than human judges. More importantly, machines given the proper databases can determine the actual likelihood that drugs will be discovered in a car when it is provided with the officer's observations.

Further, machines can overcome prejudices much more quickly than humans. Unlike humans, they can be easily coded to ignore race. Machines are obviously not a panacea for discriminatory policing. Officers may ask for the device's assistance disproportionately with minority motorists. The existence of such a device, however,

would improve even this concern. Such a machine would record racial patterns of the requests made of them, providing a valuable tool in identifying and rooting out biases in the training of police.

Using computers and metadata to evaluate probable cause and reasonable suspicion standards does not displace humanity from judging in a way that would be of concern in other multi-variable tests. Many multi-variable standards in constitutional law, for instance, similarly lack predictability. Suspicion, however, is not a matter of values or philosophy – it is truly a question of odds. There may be legitimate disagreements about the degree of certainty required to justify a search, but probable cause is stated in terms of the likelihood of success, though courts have thus far rejected efforts put actual numbers to the test. That reluctance, however, has doubtlessly been driven by the absence of any meaningful way to actually identify the odds of a search's success. [4] Machine learning thus holds the potential not only to make suspicion assessments more reliable and accurate, but to change the question courts ask in determining whether the legal standard has been satisfied. In the context of drug interdiction stops, artificial intelligence provides an opportunity to accurately assess the prediction the law required officers to make.

3. Developing a Prototype to Assess Suspicion

The current prototype still under development by Crivella Technologies, a Probable Cause Advisor, uses existing case law to identify salient features of suspicion, or stated another way, categories of suspicious facts. Examples in a drug interdiction stop include, for instance, nervousness, masking odors, and inconsistent travel plans. Artificial intelligence applications and language analysis is then used to identify various ways courts and officers describe the presence, and degree, of these salient features. A smell of talcum powder can be faint or strong; the inconsistencies of the travel plans described by the driver and passenger can be minor or extreme. And these salient features can be described in any number of ways.

Language used by a court or officer is identified as a salient factor by comparing text obtained from officers in the field against more formal language used in court proceedings and orders. Measuring semantic text similarity and the use of artificial intelligence for key decision support has been a very active area of research and development to support litigators and experts in complex mass tort pharmaceutical and medical device litigation. [5] The prototype builds upon the methods developed in these contexts and research related to Twitter tweet searches [6] and paraphrase recognition[7] to establish language translation between end user vernaculars.

Our general approach has been a vector regression model to combine a large number of textual and metadata general and domain specific features. The prototypes have been completed using proprietary marker sets tested against large litigation language corpora, in accordance with Content of Interest seed set sampling and evaluation and a powerful semantic word similarity model based on latent semantic analysis.

At present, we have assembled two major corpora of pertinent documents and language. The first, the Corpus of Judicial Probable Cause Opinions, contains over one-hundred thousand decisions assessing whether reasonable suspicion exists in drug interdiction stops. A language corpus, comprising the complete set of written words pertaining to these courts opinions, has been developed and made statistically

analyzable. Focused derivative sub-corpora have been further developed to aid in advice algorithm design, testing and training.

The second major corpora is the Language of Active Interdiction. As the name implies this corpus contains the language of police officer real time recording of interdiction observations. This corpora will identify the language officers use in the field to identify the concepts relevant to courts. In some jurisdictions, internal department requirements or judicial decrees require officers to explain their decisions to stop, detain, and search any automobile, indicating whether or not drugs were discovered. For databases of police reports identifying circumstances that led both to the discovery of drugs as well as fruitless searches, training can begin to start developing algorithms for the presence of drugs.

Databases including false-positives are somewhat obviously less available than judicial decision admitting or excluding evidence from drug interdiction stops. Every federal and state court is required to determine whether the evidence it considers in a criminal case was obtained legally. By contrast, only some jurisdictions are retaining records of the suspicion that led an officer to conduct a fruitless search. Additionally, there is often no indication in these records when officers recorded their suspicions before or after the search. Post-search statements may be tainted by facts discovered in the search. Nevertheless, the limited records that exist provide training material so the system can learn the various ways salient factors are described and, more substantially, begins the development of an algorithm for predicting whether drugs are actually present – as opposed to merely the odds that a court will find the search justified. The system will obviously be able to predict judicial behavior with far greater accuracy than the presence of drugs in its early phase.

The trial implementation phase of the prototype will, however, begin the process of producing the best version of a database of suspicion. There will be substantial law enforcement benefits to using even early pilot versions of the program that predict only the judicial perspective on the officer's observations. Officers who type the facts they observe into the program prior to a search, or continued detention, will discover the odds that a court would find a detention or search justified, enhancing their efforts to justify the search to a judge hearing a motion to suppress if drugs are discovered. In entering the data prior to search, the officers will be building a system that increasingly is able to conclude just how probable the officer's cause is.

References

[1] *Warrantless Searches and Seizure*, Geo. L. J. 91 (2003), 36-54.

[2] O.W. Holmes, Jr., *The Path of the Law*, 10 Harv. L. Rev. (1897) 457-478.

[3] D.H. Kaye, *The Double Helix and the Law of Evidence* (2010).

[4] B. Alarie, A. Niblett & A.H. Yoon, *How Artificial Intelligence Will Affect the Practice of Law*, Univ. of Toronto L. J. 68 (2018) 106-24.

[5] B. Siriam, D. Fuhry, E. Demir, H. Ferhatosmanoglu, M. Demirbas, *Short Text Classification in Twitter to Improve Information Filtering*, Proceedings of the 33rd Int'l AEM SIGR Conference in Research Development in Information Retrieval, 841-42.

[6] B. Dolan, C. Quirk, C. Brockett, *Unsupervised Construction of Paraphrase Corpora: Exploiting Massively Parallel News Sources*, Proceedings of the 20th Int'l Conference on Computational Linguistics, Article No. 350 (2004).