

RESEARCH

Open Access

Distribution of event complexity in the British Columbia court system an analysis based on the CourBC analytical system

Amir H Ghaseminejad*, Paul Brantingham and Patricia Brantingham

Abstract

This paper reports an exploratory research on the distribution of event complexity in the British Columbia court system. Analysis of event distribution shows that the frequency of events sharply decreases with the increase in the number of persons and counts. The most frequently observed type of event is the event that has one person involved with one count. The number of events observed sharply declines when we query for events with a larger number of people involved or more counts charged. It is found that the number of events observed exponentially decreases when more complex events comprising more counts are analyzed. The same exponential decrease is observed for events with two or more people. This means that, in general, the least complex events are the most frequently observed ones. The events with more than one person involved have a mode of two counts. A first approximation model for the distribution of the load on the system based on different levels of complexity is proposed. The proposed model can be used for and be evaluated by predicting the load distribution in the BC criminal court system.

Keywords: Event complexity, Counts, Courts, Persons, Distribution, Data mining, Entity relationships, Charge, Event

Introduction

Common law criminal justice systems have experienced a series of major problems in recent years including declining rates of case clearances and prosecutions, rising rates of remand in custody, increasing delays between charge and trial dates, and increasing rates of case collapse at trial [1] [2] [3] [4]. Sources of these problems may reside at many points in the system with important operational feedbacks between decisions and events occurring in the investigative, prosecutorial, judicial and correctional elements of the system [5]. Case complexity is thought to be a major contributor to justice system problems, but little systemic science addresses the issue at any agency level [6] [4].

Interoperable justice agency databases could be used to identify systemic trouble points relate them to case complexity and perhaps develop improvements. For a variety of good reasons ranging from issues of privacy to the requirements of fair trials judicial data systems are

rarely made available to researchers interested in understanding system problems.

The Institute for Canadian Urban Research Studies (ICURS) has collected information about all crimes handled in the British Columbia criminal court system during a 3-year period encompassing 2008 through 2010 using publically available court data published by the courts from a judicial Records Management System named JUSTIN. These published data have been reverse engineered into a research data warehouse called CourBC. A research tool called the CourBC Analytical System has been developed to facilitate repetitive queries and data mining in this database.

A file, in a British Columbian criminal court, represents all linked courtroom actions. This would involve crimes committed by one or more individuals, that is, all crimes associated with co-accused. The crimes may be divided into informations or indictments when the crimes are linked by occurrence time or place, called folders in CourBC. A file is processed in a specific court and has a folder number, one or more associated documents, one or more persons involved, one or more crime codes involved,

* Correspondence: ahghasem@sfu.ca
Institute for Canadian Urban Research Studies, Simon Fraser University,
Vancouver, Canada

and one or more counts arising from those crime codes [7]. Figure 1 is an Entity Relationship Diagram showing the structure of the relationship between entities in the court system. A unique folder in a court can contain many documents, a document in a court-folder can be associated to many persons and a person can be associated to many documents in a folder in a court. In this paper, we refer to a unique folder in a court as an event. A document associated to a person can be related to many crime codes. Moreover, one or more counts can be assigned to the same crime related to a person and a document.

One of the capabilities of the CourBC Analytical System is to find the number of unique court events based on the attributes of the file in court system.

Event complexity

The data in the CourBC Analytical System includes court name and location, hearing date, file number, folder number which uniquely identifies an event in a court, document number, number of counts; as well as bail and custody status, crime statute, section, subsection, next appearance detail, findings, results and convictions. It also includes the scheduling information which enables the researchers to track an event as it goes through hearings and identify the result for each count. Scheduling data includes reason for appearance, hearing time in the day, information about plea, age of the file in the court system measured in number of days. This enables the researchers to analyze the events in terms of the number of people involved, the severity of crime, and the type of the verdict. On the one hand, the CourBC dataset enables the researchers to conduct descriptive studies on the distribution of crime types and to analyze the differences in varying court locations; on the other hand, this data makes it

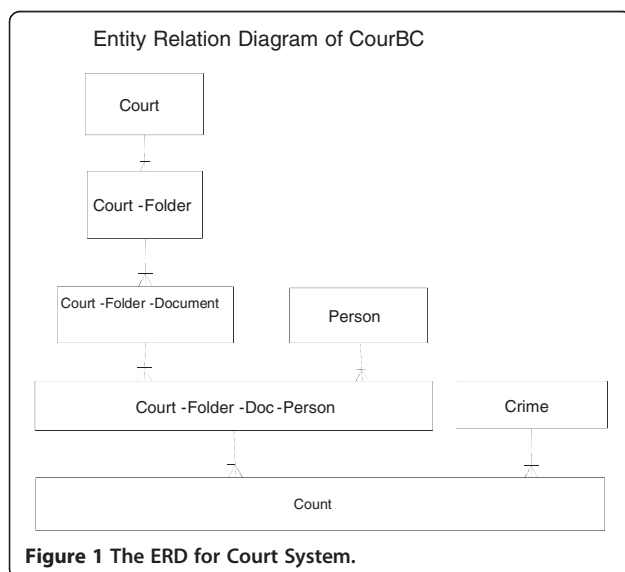
possible to study the impact of crime complexity and other determinants on the number of hearings for each event in the court system. Moreover, the data has the potential to be used for modeling the central tendencies and the distribution of outcomes as a dependent variable influenced by variables such as crime severity, number of people involved, and location.

Events are handled in one or more court appearances. An event may be simple, that is, have clear evidence and present no difficult legal issues. An event, on the other hand, may involve many accused with multiple defence counsel, include a large number of charges for specific crimes, involve multiple witnesses, and involve legally untested issues or any combination of these complicating attributes. Complexity is the term used to describe the differences between the simple and straightforward case and one with many distinct and inter-related components.

The complexity of an event is a determinant of the police, legal system and correctional resources that the event will probably consume. Event complexity is a complicated construct and is a predictor for the overall load imposed by an event on the overall criminal justice system [5].

The essential qualities of event complexity are the type of the crime, the number of people involved in the crime, the number of persons involved, the evidence collected, and the event's legal and moral severity [6] [8] [9]. The court system can be modeled as a relational database that identifies criminal events as folders in each court. Each folder in a court may include one or many documents. One document can be related to many persons and a person can be charged for many counts of the same crime.

We can theorize that event complexity has a positive correlation with number of persons, number of documents in the folder, number of counts charged and the type of crime [1] [6] [8].



$$EC = f(p, d, c, ct) \quad (1)$$

Where:

EC is event complexity

p is number persons

d is number of documents

c is number of counts charged and

ct is crime type.

All criminal cases, and therefore all events tracked in COURBC, are marked by various legal issues having varying levels of legal uncertainty. This level of legal uncertainty is a determinant of event complexity but can only be assessed through a review of the facts of the case. The likelihood of legal uncertainty increases with the number and novelty of legal issues presented. The number of legal issues presented is likely to increase

with the number of persons charged, the number and severity of crimes charged, and the severity of potential penalties upon conviction. In addition, the human agents involved in the event, (lawyers, judges, prosecutors) and their backgrounds and history can influence the complexity of the case. This may lead to events that, from “person, document, and count” point of view, are simplistic but still have a certain internal complexity [6]. CourBC cannot, at present, address these components of complexity, but such legal and agent based complexity is relatively rare as indicated by the limited number of cases appealed to higher courts.

In this paper we present a first approximation model for the complexity of events that cover a great majority of events.

Findings

In this section we present our analysis of the characteristics of the data, and will highlight the distribution of event as it was observed. As is shown in Figure 2, Figure 3, and Figure 4, the conditional probability of having more counts with a chosen number of docs has a positive correlation with the number of documents in the file.

$$P(c|d) \propto d \quad (2)$$

Where:

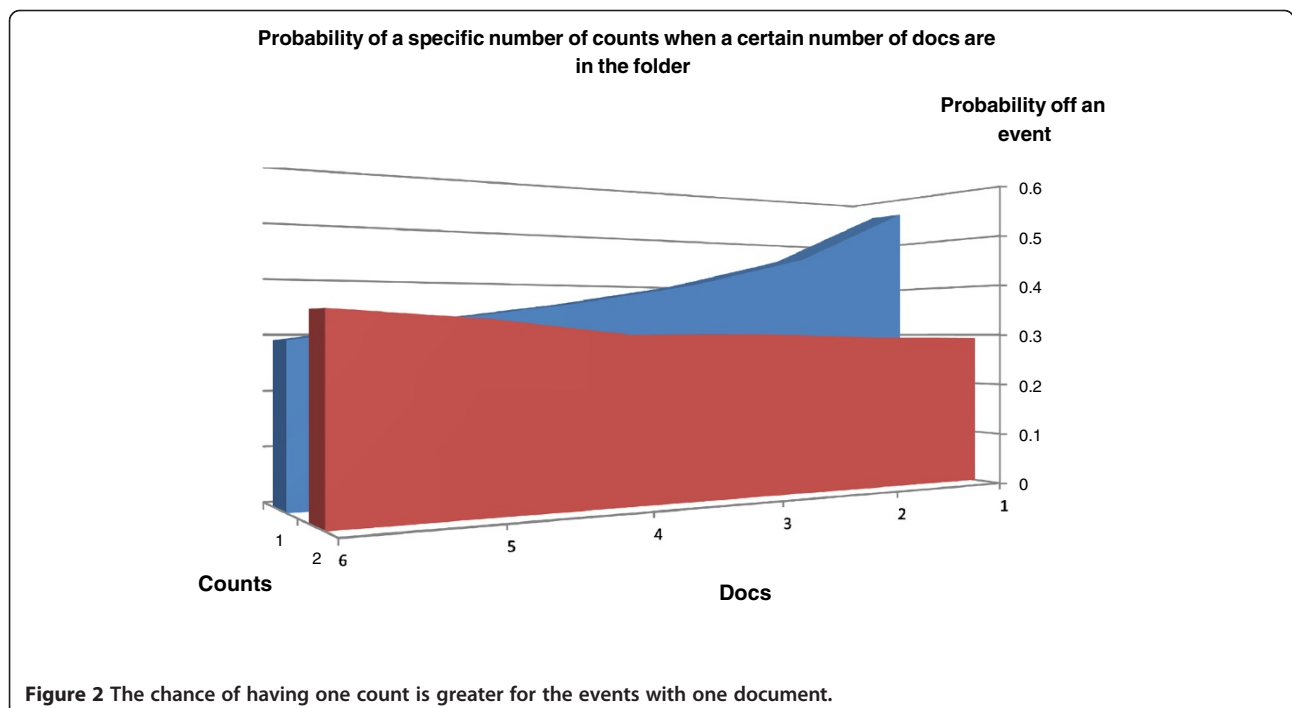
i is the maximum counts charged in the folder
 d is the number of documents in the folder.

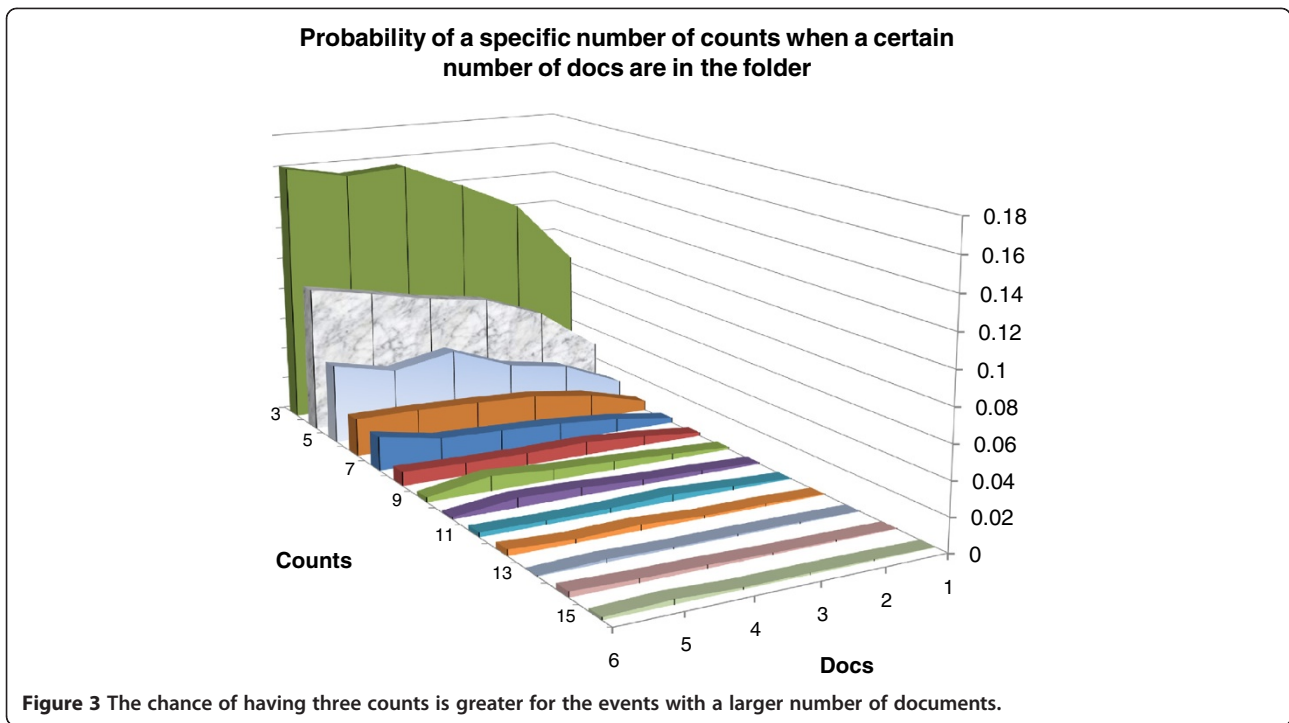
In Figure 2, it is shown that the chance of observing a single count is higher in events that have a single document and declines steadily as the number of documents in the event increases.

On the other hand, in Figure 3, the chance of having three counts is greater for the events with a higher number of documents. Similarly, the chance of comprising five counts is greater for events with six docs than the events with one document.

Figure 4 shows that the chance of an event with 14 counts is higher for events with six docs than events with only one document. Therefore, we can conclude that the number of documents and the number of counts are not independent variables. Counts in the folder are known to be a good predictor of complexity of the events because of the decisions that must be made about all of them during the course of prosecution. In this paper, we will assume that counts are a good proxy variable for the influence of the number of docs on complexity.

In this paper, consistent with past studies of case complexity and case processing [8], complexity is considered as a combination of two variables: counts and persons. The frequency distribution of events in terms of these two variables is analyzed. We study a cross section of the events being processed in the court system over a 3 year period. Each event in our dataset is analyzed to identify the number of persons involved in the case, the maximum doc number associated to the event, and maximum number of counts assigned to the events.





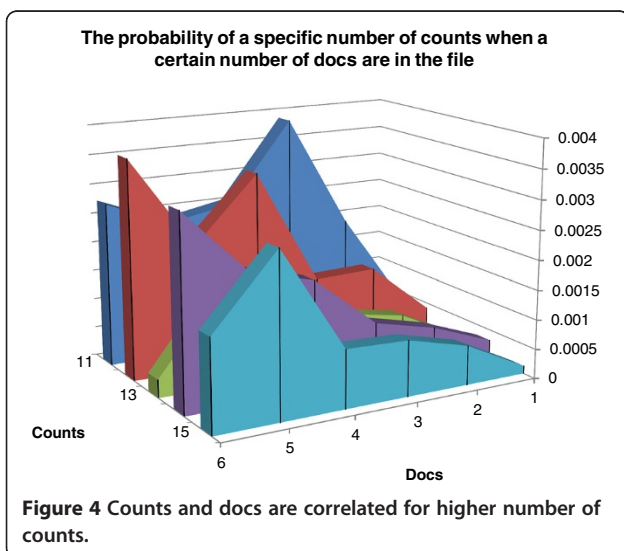
There are events that are processed within the time interval of our cross section but have been brought to court before our time interval. For some of these, events had hearing(s) before our time window and some of the docs or persons or counts may have been eliminated (legally concluded) from further hearings. Therefore, we expect the complexity of these events may have been more than what we observe and overall complexity of the events being processed in the system may be slightly greater than what we observe. All that

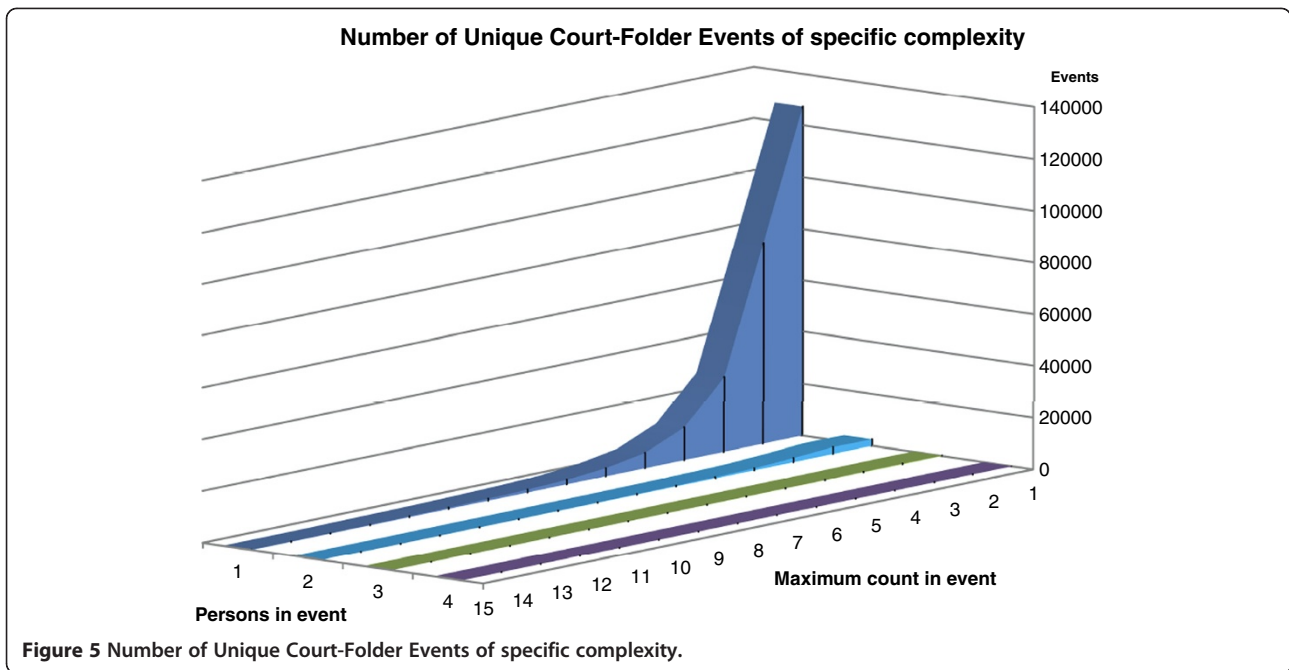
considered, the limitations do not impede the significance of our finding about the distribution of complexity in our cross section time window. In other words, our observation is a plausible description of the complexity of the events as they were being processed in the time window.

Our study shows that the distribution of events is such that, regardless of number of docs in the folder, the frequency of events sharply decreases with the increase in the number of persons and counts. As shown in Figure 5, the most frequently observed type of event has one person involved with one count. The number of events observed sharply declines when we query for events with a larger number of people involved or more counts charged.

Figure 6. Shows how even for events with two people involved the number of events observed exponentially decreases when the event complexity increases.

This means that in general the least complex events are the most frequently observed ones. However, we found that the cases with more than one person involved have a mode that is at two counts. We turn in the next section to development of a model of case complexity and court workloads derived from these findings that we think could be used to help identify and address both existing and potential case handling trouble point. Linking this model to models derived from a compatibly defined police information database may allow us to begin to identify dysfunctional feedbacks between these two components of the criminal justice system.





Abstraction from findings

In this section, we propose a first approximation linear model for the distribution of case complexity and claim that the distribution of the load on the justice system can be predicted based on this model. In this model, event complexity has a positive correlation with counts and a positive correlation with persons in the file.

$$EC \propto k * p * c \quad (3)$$

Where:

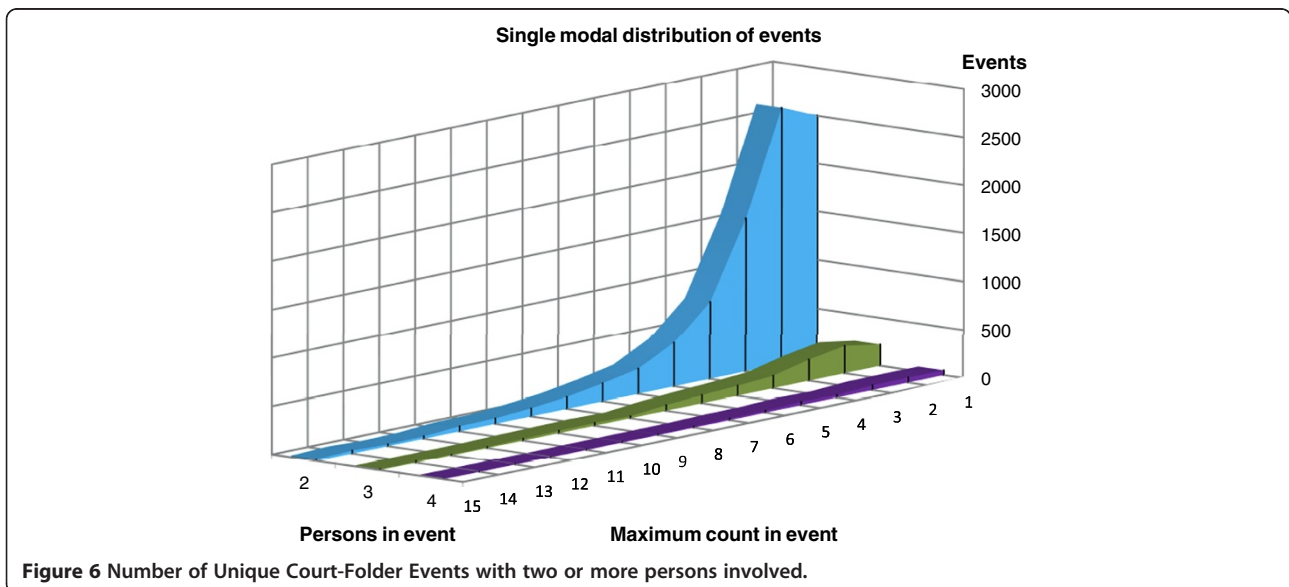
EC is event complexity

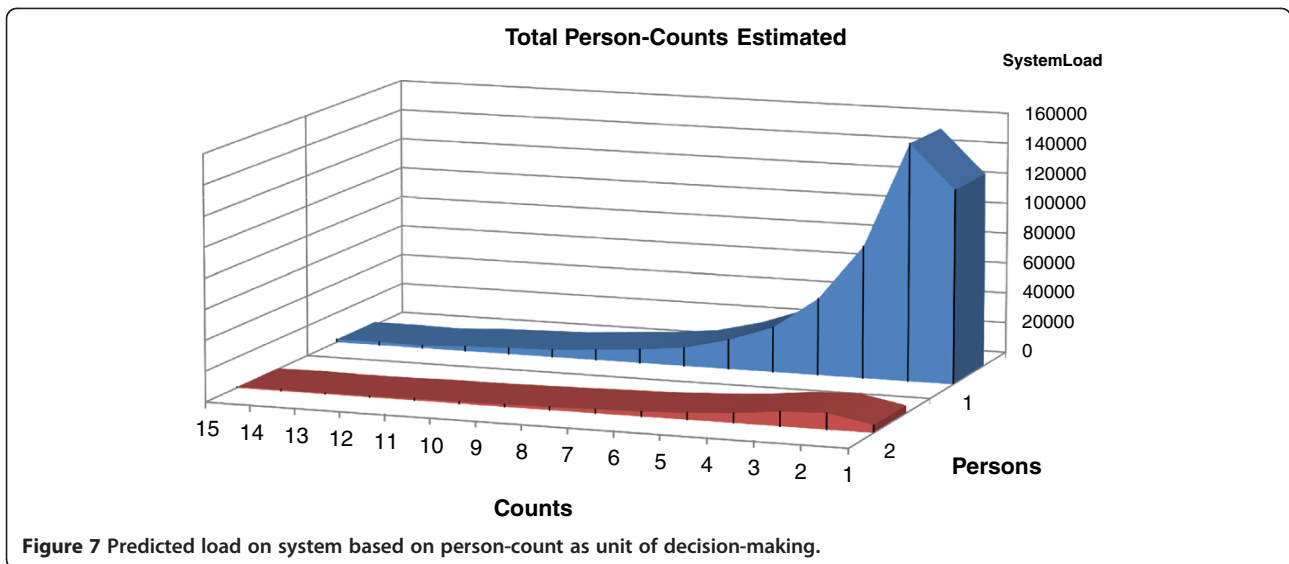
p is the number of persons in the folder
 c is the maximum counts charged in the folder.

k is a coefficient that presents the sensitivity of event complexity to number of persons and maximum counts in the folder.

As shown in the relationship shown above, event complexity grows proportional to the number of persons in the folder and the maximum counts charged in the folder.

We can operationalize event complexity by measuring the load on the justice system. One probable approach to measuring the load on the system is analyzing the number





of hearings occurring in different types of events. Our Hypothesis is that the number of hearings related to different events is proportional to the number of persons and counts in the event. The following formula shows the first approximation linear model of this relationship.

$$LS \propto k * f * p * c \quad (4)$$

Where:

LS is load on the justice system

f is the frequency of a folder with a specific person-count being observed

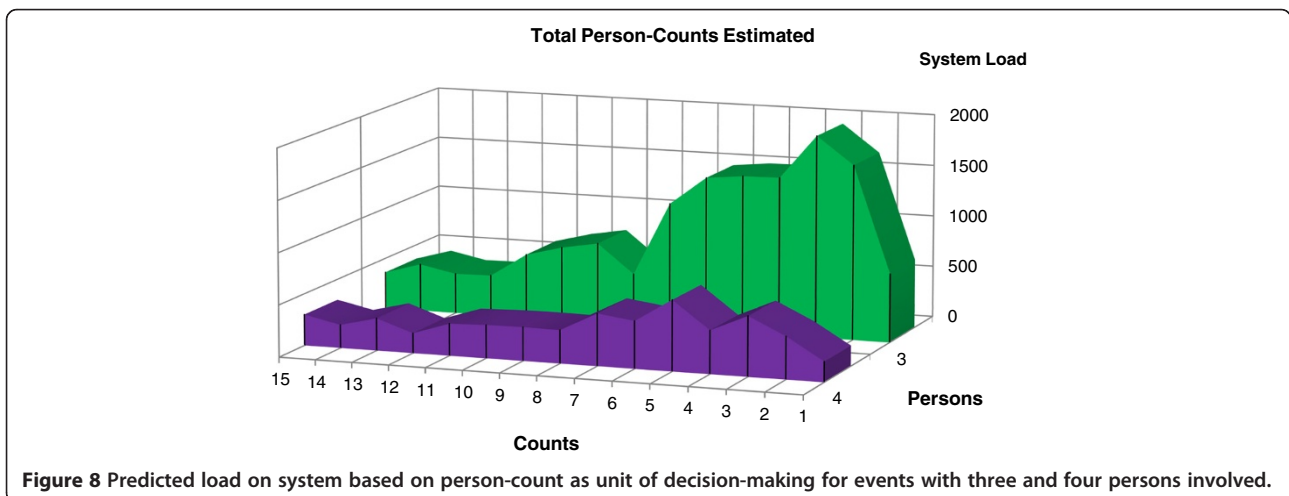
p is the number of persons in the folder

c is the maximum counts charged in the folder

k is a coefficient that presents the sensitivity of the load on the system to the complexity

In the relationship shown above, load on the justice system grows proportional to frequency of observing a folder with a specific person-count being observed and the number of persons in the folder and the maximum counts charged in the folder.

Based on this model, as shown in Figure 7, we have calculated the estimated load on the system measured with the number of person-counts that should be decided by the court system. We can hypothesize that the load on the system for the most frequently observed events involving one person should have a mode on events with two counts. Also, the load related to events with two persons involved, has a mode of two counts. However, as shown in Figure 8, events with three persons involved have a mode at three counts. Similarly, the mode for four persons involved is observed at a higher number of counts.



As it is clear from these figures, we expect the share of complex events in the load on the system grow nonlinearly. The reason why our proposed multi variable linear models have produced the non-linear results is in the distribution of the real events. The frequency of the events with higher complexity is non-linearly less than the frequency of simpler events. According to our data the number of people and counts in complex cases do not compensate to bring the overall load of complex events the prominence of the share of simpler cases in system's load. We are going to evaluate this model by comparing the number of hearings of the events with the predictions of our model and verify if the assumption of linear relationship will survive, or we may need to add a function to account for the relationship of the number hearing and event attributes to our model.

Conclusion and future research

The observations reported in this study set the foundations for modeling the distribution of event complexity. This distribution can be used in both process modeling, and agent based modeling of the criminal justice system where there is a need to test the model with cases of different complexities.

We are planning to find the best-fit distribution to our crime complexity data in a parsimonious way. Such a distribution must be plausible from the theoretical perspective and must be evaluated with a reasonable goodness of fit. The possible variables should be analyzed for co-linearity and only the important independent ones should be included. We expect that such analysis will be confirmatory and consisting with the proposed model in this paper.

To test the hypothesis proposed in this paper predicting the load on the system, we will analyze the number of court hearings/appearances for all events under the classification used in this paper. If the predictions of this paper are true we expect to observe more hearings for cases with more people and more counts involved but we expect that most number of hearings involve events with one person and two counts. Similarly, among events with two persons events with two counts are expected to dominate court workloads. While the modal point may be two, it will be important to see how rapidly the number of court hearings/appearances increases as complexity increases for higher numbers of persons and counts. It is possible to observe that the load on the system from more complex events is more than their proportional frequency. If such thing is observed it means that complex events' share of the load on the system grows nonlinearly. We will attempt to develop a multi variable linear (if necessary non-linear) model of the relationship between load on the system and the dimensions of crime complexity. Such a model will enable the

decision makers to predict the expected load on the system as result of events based on their complexity. Moreover, we are planning to add crime type and crime severity into our complexity model in the next round of analysis. There is good reason to think that different types of crime present different levels of complexity in court. In addition, a large number of legal system and law enforcement analyses could be performed if we could have access to court documents and law enforcement records in database form and for a wider time frame. We think that we will be able to link this database to police databases so that we will be able to test the proposition that the complexity of cases in court is grounded in the complexity of cases as they present to police.

Once COURBC is linked to police databases, we plan to explore the overall extent to which case complexity at different points in the criminal justice system drives the overall use of resources by the system.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

PLB conceived and developed the original project. AHG developed the original data collection tool. AHG, PLB and PJB supervised data collection and collaborated in development of the underlying database. AHG developed the analytic system and conducted the analysis. PLB and PJB consulted on analysis and interpretation of results. AHG, PJB, PLB co-wrote the manuscript. All authors read and approved the final manuscript.

Received: 31 August 2011 Accepted: 15 May 2012

Published: 22 August 2012

References

1. M Gannon, M Karen, B Karen, T-B Andrea, K Rebecca, *Criminal Justice Indicators* (Statistics Canada, Ottawa, 2005). Catalogue No. 85-227-XIE
2. Statistics New Zealand, *Crime in New Zealand: 1996-2005* (Statistics New Zealand, Wellington, 2006)
3. J Thomas, *Adult Criminal Court Statistics 2008-2009. Juristat 2010*, 30, No. 2 (Canadian Centre for Justice Statistics, Ottawa, 2010). Statistics Canada Catalogue Number 85-002-X
4. W Michael, Inmate Perspectives on the Remand Crisis in Canada. *Canadian Journal of Criminology and Criminal Justice* 51, 355-379 (2009)
5. A Azadeh, B Patricia, B Paul, D-A Vahid, F Ron, F Ellen, H Ghaseminejad Amir, G Christopher, L Jenny, P Nahanni, R Alexander, A van der Waall, Utilizing Simulation Modeling for Criminal Justice System Analysis, in *Artificial Crime Analysis*, ed. by L Lin, E John (Idea Group Publishing, Hershey, PA, 2008)
6. M Heise, Criminal Case Complexity: An Empirical Assessment. *J Empir Leg Stud* 1, 331-369 (2004)
7. J Jordan, *The Administrative and Information Technology Environment of the Superior Courts Judiciary in British Columbia* (AIJA Law and Technology Conference, Sydney, Australia, 2008)
8. Dg Horne, *Case complexity and case processing time* (unpublished ma thesis 1981, Simon Fraser University, Burnaby, b.c, 1981)
9. M Schlanger, *Comment on Measuring Case Complexity*, Empirical Legal Studies Blog posted September 25, 2006 www.elsblog.org. Accessed March 12, 2011

doi:10.1186/2190-8532-1-13

Cite this article as: Ghaseminejad et al.: Distribution of event complexity in the British Columbia court system an analysis based on the CourBC analytical system. *Security Informatics* 2012 1:13.