A Simple Way to Maximize the Number of Hits From Ballistic Imaging When Processing Capacity is Limited

By: Can Wang, M.S.[*], Mardy Beggs-Cassin, M.S.F.[†], Lawrence M. Wein, Ph.D.[‡]

Keywords: ballistic imaging, forensic science, mathematical modeling, operations management

ABSTRACT

Many U.S. organizations lack the resources to enter all cartridge cases from crime scenes and test-fires into the National Integrated Ballistic Information Network and search for hits. This paper summarizes the main question in (1): if ballistic imaging capacity is limited (i.e., less than the arrival rate of newly acquired cartridge cases), which specific cartridge case types should be processed to maximize the number of cartridge cases that generate at least one hit? This problem is solved using a mathematical model, and data from Stockton, CA, are used to show that a simple scheme– prioritize crime scene evidence over test-fires, rank calibers by their likelihood of achieving a hit, and process only the higher-ranking calibers– can significantly increase the number of hits generated.

Introduction

The National Integrated Ballistic Information Network (NIBIN) maintains a national database of images of cartridge cases that have been either recovered as crime scene evidence or test-fired from confiscated guns. This infrastructure allows U.S. organizations to use computerized ballistic imaging technology to generate hits, which either link crimes or reveal a cold connection between a confiscated gun and a past crime. However, there is immense intercity variability in the use of NIBIN [2, 3, 4], and an estimated 72% of hits in the U.S. come from the 20% of cities that enter the most cartridge cases into NIBIN [2].

This paper presents the main results of "Optimizing Ballistic Imaging Operations," which contains a long technical analysis and is referred to here as WBW [1], in a concise, non-mathematical manner. WBW attempts to gain a better understanding of ballistic imaging operations by using several years of data from Stockton, CA, and a mathematical model to address two issues: (i) how the number of hits vary with the proportion of cartridge cases that are processed (i.e., entered into NIBIN and compared to database entries in the search for hits), and (ii) if there is not sufficient capacity to process all cartridge cases, which cartridge cases should be processed to maximize the number of hits. In the process of addressing (ii), a simple prioritization scheme is proposed, which has the potential to significantly increase the number of hits generated by a city with limited processing capacity.

Materials and Methods

A brief overview of the methods is provided here, and readers are referred to WBW for details.

Date Received: April 14, 2017
Primary Review Completed: August 29, 2017
Secondary Review Completed: March 25, 2018

Data

Ballistic imaging in Stockton, CA, which was the second most violent city in CA in 2012 [5], underwent major changes in personnel and processes in 2012-13. They stopped relying on a state crime laboratory, and used firearms technicians (including co-author Mardy Beggs-Cassin) to enter newly-acquired cartridge cases into NIBIN and identify high-confidence candidate hits, and hired a part-time contract examiner to make the final determination [5]. These changes allowed them to start processing all newly-acquired cartridge cases and to subsequently generate a three-fold increase in hits [5].

Two data files from Stockton were used. The first file contained information about 6703 cartridge cases that were entered into NIBIN between January 1, 2010, and March 3, 2015. Each NIBIN entry is characterized as one of six types: evidence from homicides, assaults with a deadly weapon, and other crimes, as well as test-fires from firearms involved in homicides, assaults with a deadly weapon, and other crimes. Each NIBIN entry was additionally characterized as one of 12 calibers (see **Table 1**). The second data file contained information about 964 high-confidence hits detected from July 26, 2011, to February 20, 2013, and from June 18, 2013, to March 3, 2015 (data pertaining to four months of hits are missing). For each hit, the entry type and caliber of both cartridge cases in the match are specified. Some cartridge cases generated more than one database match in this data file.

[*]Electrical Engineering Department, Stanford University, Stanford, CA 94305, canw@stanford.edu

[†]Firearms Unit, Stockton Police Department, Stockton, CA 95202, Mardy.Beggs@stocktonca.gov

[‡]Graduate School of Business, Stanford University, Stanford, CA 94305, lwein@stanford.edu, corresponding author

| Crime Type (*i*)	Description	Caliber (*j*)	Description	Caliber (*j*)	Description
1	homicide	1	12 gauge	7	38 special
2	assault with deadly weapon	2	223 Remington	8	40 Smith & Wesson
3	other	3	22 long rifle	9	45 auto
4	test-fire: homicide	4	25 auto	10	7.62 × 39 mm rifle
5	test-fire: assault with deadly weapon	5	357 Magnum	11	9mm Luger
6	test-fire: other	6	380 auto	12	other

Table 1: The entry types *i* = 1, ... , 6 and the calibers *j* = 1, ... , 12 taken from Table 1 of WBW

Parameter Estimation

The mathematical model has several types of parameters that are estimated in WBW using the Stockton data. The total arrival rate of cartridge cases was estimated to be 4.10 per day, which did not vary significantly from year to year. The proportion of all arriving cartridge cases that were of a specific entry type and specific caliber was also required for the analysis. Because there are six entry types and 12 calibers, hmice were 72 of these values. Also required was the proportion of all arriving cartridges of a specific entry type and caliber that resulted in at least one hit, generating 72 values that we refer to as "hit probabilities." The average number of hits for each arriving cartridge case that generated at least one hit is estimated to be 1.92.

Hit Analysis

The key performance metric is the "hit proportion" (not to be confused with the 72 hit probabilities defined above), which is the proportion of all arriving cartridge cases that generate at least one hit. Multiplying the hit proportion by the total arrival rate of 4.10 cartridge cases per day gives the number of cartridge cases per day that generate at least one hit. Hence, assuming processing capacity is the limiting factor, maximizing the hit proportion is equivalent to maximizing the number of cartridge cases that generate at least one hit over any given time period. The hit proportion depends on the rate at which cartridge cases of each entry type/caliber pair is processed, which in turn depends on the total amount of processing capacity (we define processing capacity as the number of cartridge cases that are processed per day) and how that capacity is allocated among the 72 entry type/caliber pairs. We are primarily concerned with the case in which processing capacity is less than or equal to the arrival rate. After all, if the processing capacity was greater than the arrival rate (which in Stockton's case was 4.10 cartridge
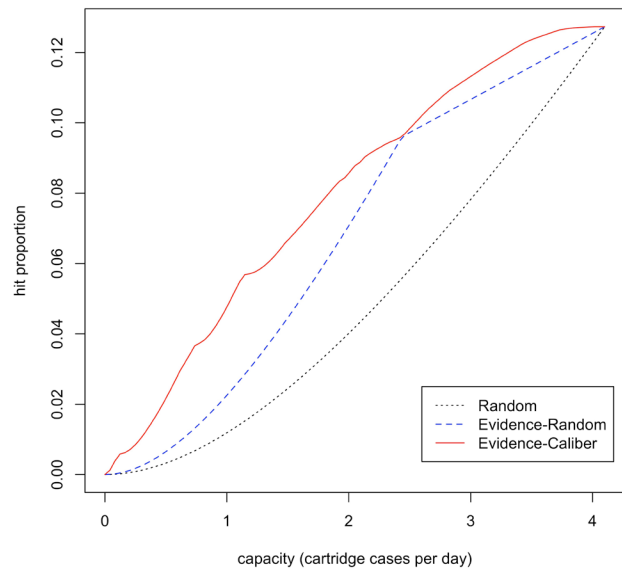
Figure 1: The hit proportion results for the three capacity allocation policies described

cases per day), then all arrivals could be processed and there would be no need to allocate capacity.

WBW derives a mathematical expression for the hit proportion in terms of any possible capacity level (between 0 and the arrival rate of 4.10 cartridge cases per day) along with any allocation of this capacity among the 72 entry type/caliber pairs, and in particular finds the hit proportion for nine policies that are specified in Table 3 of WBW; three of these nine policies are described here. In addition, for every possible capacity level, WBW also derives the optimal allocation of this capacity among the 72 entry type/caliber pairs.

Results

The results appear in **Figure 1**. The horizontal axis displays how many cartridge cases per day the city is capable of processing, varying from zero to 4.10 cartridge cases per day, which is Stockton's arrival rate. The vertical axis displays the proportion of all arriving cartridge cases that generate at least one hit. This relationship between capacity and the hit proportion depends on the allocation policy, i.e., which of the 72 entry type/caliber pairs are processed when processing capacity is less than 4.10 cartridge cases per day. All policies have a hit proportion of zero when the capacity is zero and have a hit proportion of 0.127 – which is the hit proportion achieved by Stockton when it processed all of its cartridge cases – when the capacity equals 4.10 cartridge cases per day. Hence, each allocation policy generates a curve in **Figure 1** that starts at the point (0, 0) and ends at the point (4.10,

0.127), and higher curves correspond to better allocation policies. Note that the shapes of the curves in **Figure 1** do not depend on the value of the arrival rate of cartridge cases: regardless of the value of the arrival rate, the left portions of the curves represent scenarios in which only a smaller proportion of the arriving cartridge cases can be processed, and the right portions of the curves represent scenarios in which a larger proportion of the arriving cartridge cases can be processed.

Figure 1 describes the performance of three allocation policies, with the lowest curve corresponding to the Random Policy, which represents any policy that does not use entry type/caliber information to allocate capacity. Two examples of the Random Policy include maintaining a queue of all yet-to-be processed cartridge cases and choosing the next cartridge case to be processed by either selecting one at random or using First-Come First-Served (e.g., as is used in a typical service system, such as a supermarket or a fast-food restaurant).

To gain intuition about the performance of the Random Policy, assume for the moment that there is at most one match in the database for any arriving cartridge case. Now suppose that the processing capacity is exactly one-half of the arrival rate, which is 0.5 x 4.10 = 2.05 cartridge cases per day. Suppose an arriving cartridge case has a mate that arrived at some earlier time. To obtain this hit between the arriving cartridge case and its mate, two things must happen: the current arrival must be processed, which has a 50-50 chance of occurring (because only half of arriving cartridge cases are processed), and the mate had to have been processed at the (earlier) time of its arrival, which also has a 50-50 chance. Hence, the likelihood of the current arrival generating a hit is 25%, the same as the likelihood of getting two heads from two consecutive coin tosses. Therefore, the hit proportion for the Random Policy when the processing capacity is 0.5 x 4.10 = 2.05 cartridge cases per day equals 0.25 times the potential hit proportion of 0.127, which is 0.25 x 0.127 = 0.03175 cartridge cases per day.

By the same argument, if the processing capacity is 10% of the arrival rate (i.e., 0.1 x 4.10 = 0.41 cartridge cases per day) then the hit proportion is only 1% of the potential hit proportion (i.e., 0.01 x 0.127 = 0.00127 cartridge cases per day). This quadratic performance, where the hit proportion increases with the square of the ratio of the capacity and the arrival rate (i.e., equals *(capacity ÷ arrival rate)² x 0.127*), is in stark contrast to virtually every manufacturing (e.g., auto plant) or service (e.g., fast-food restaurant) operation, where the output increases linearly (i.e., increases with *(capacity ÷ arrival rate)*). This is because in a typical operation, every

item that is processed corresponds to salable output, whereas in ballistic imaging, the output corresponds to a hit, which requires processing an arrival and its mate in the database.

Recall that the thought experiment above assumes no arrivals have more than one hit in the database. However, in Stockton, each arriving cartridge case that generates at least one hit has on average 1.92 hits. Hence, the lowest curve in **Figure 1**, while still convex (and hence below the straight line that connects the endpoints (0, 0) and (4.10, 0.127), is slightly higher than *(capacity ÷ arrival rate)² x 0.127*.

The next policy is the Evidence-Random Policy, which gives priority to all crime scene evidence over all test-fires (i.e., test fires are only processed after all evidence is processed), but otherwise chooses randomly among evidence and among test-fires. In Stockton, evidence comprises approximately 60% of all arrivals and 80% of all hits. Hence, the performance curve generated by this policy consists of two convex curves, with the left segment for evidence and the right segment for test-fires. This policy performs well when capacity is approximately 60% of the arrival rate (i.e., 0.6 x 4.10 = 2.46 cartridge cases per day), but otherwise can incur significant suboptimality.

The third policy is referred to as the Evidence-Caliber Policy, and its strong performance (it has the highest curve in **Figure 1**) is the main result of this study. As with the Evidence-Random Policy, this policy gives all evidence priority over all test-fires. However, within evidence, cartridge cases are grouped by caliber. This grouping changes the left quadratic curve of the Evidence-Random policy into 12 smaller convex curves that are visible upon close inspection of **Figure 1**. These 12 calibers are ranked according to their hit probability (recall that the Stockton data allows for the estimation of the hit probability of each of the 72 entry type/caliber pairs and higher-ranking calibers are given higher priority. Within evidence of the same caliber, cartridge cases are prioritized (among homicide, assault with a deadly weapon, and other crime) by their hit probabilities. Within test-fires, the 36 entry type/caliber pairs are also prioritized by their hit probability (note that test-fires do not ordinarily match each other, and so there is no need to group them by caliber).

Figure 1 shows that the Evidence-Caliber Policy performs much better than the Random Policy: e.g., it more than doubles the number of hits when only half of the arriving cartridge cases are processed. Figure 2 of WBW shows that the Evidence-Caliber Policy performs nearly as well as the optimal policy, which is numerically computed in WBW.

Prior to 2013, Stockton not only gave priority to evidence

over test-fires, but also gave priority to (evidence and test-fire) homicides (which has an arrival rate of 0.60 cartridge cases per day, or 14.6% of all arrivals) over non-homicides. In WBW, a policy that prioritized all homicides over all non-homicides but otherwise used an optimal allocation was tested. Because homicides and non-homicides have a similar hit probability in Stockton, Figure 2 in WBW shows that the performance of this policy is similar to that of the Random Policy when the processing capacity is less than 0.60 cartridge cases per day, and steadily improved until it performs nearly as well as the Evidence-Caliber Policy when the processing capacity is greater than 1.80 cartridge cases per day (which is 43.9% of the arrival rate). That is, performance is not sacrificed by prioritizing homicides if at least 43.9% of arrivals are processed.

Operationalizing the Evidence-Caliber Policy

To aid in the possible application of these ideas, we present a five-step approach that operationalizes the Evidence-Caliber Policy. That is, for a given capacity, this procedure decides which entry type/caliber pairs are processed. To describe this approach precisely, we need to introduce mathematical notation.

Step 1: Specify the number of entry types and calibers. Specify I and J, where there are I entry types indexed by $i = 1, ... , I$ and J calibers indexed by $i = 1, ... , J$. As in WBW, let $i = 1, ... , I/2$ (i.e. the first half of the series) be evidence and $i = I/2 + 1, ... , I$ (i.e. the second half of the series) be test-fires. Although Stockton had $I = 6$ and $J = 12$ (after aggregating many rare calibers, which each represented less than 1% of all arrivals, into a single "other" category), I and J may vary by agency; e.g., if an agency's database has a finer crime classification than homicide, assault with a deadly weapon and other crimes, then $I > 6$, and if an agency's database does not classify crimes then I=2.

Step 2: Estimate the arrival proportions and the hit probabilities. For each (i, j) pair, calculate a_{ij}, which is the proportion of all historical NIBIN entries that are of a particular type of (i, j). For each (i, j) pair, also calculate the hit probability h_{ij}, which is the proportion of historical NIBIN entries of type (i, j) that had at least one hit (i.e., the number of type (i, j) entries that had at least one hit divided by the number of type (i, j) entries).

Step 3: Estimate the arrival rate and capacity. Let a be the average number of cartridge cases that arrive per day and let c be the processing capacity, i.e., the average number of cartridge cases that can be processed per day. We assume that

capacity is limiting, i.e, a is larger than c, and therefore we can process only a proportion c/a of arrivals.

Step 4: Rank the calibers. The hit probability for caliber j is $\sum_{i=1}^{I} a_{ij} h_{ij} / \sum_{i=1}^{I} a_{ij}$, which is referred to as h_j. Define the caliber rankings $(1), ... , (J)$ such that $h_{(1)} \geq h_{(2)} \geq ... \geq h_{(J)}$, so that caliber (1) is the highest-ranked caliber and caliber (J) is the lowest-ranked caliber.

Step 5: Implement the policy. There are two cases:

Case 1: The arrival rate of evidence cartridge cases, which is $a \sum_{i=1}^{I/2} \sum_{j=1}^{J} a_{ij}$ and denoted by a_e, is greater than the capacity c. In this case, we do not process any test-fires, and process only a portion of evidence. Evidence cartridge cases of caliber (j) are ranked lower than evidence cartridge cases of caliber $(j - 1)$ and ranked higher than evidence cartridge cases of caliber $(j + 1)$. Among evidence cartridge cases of caliber j, there are $I/2$ entry types (in WBW, these are homicide evidence, assault with a deadly weapon evidence, and other crimes evidence), and these $I/2$ entry types are ranked according to higher values of the hit probability h_{ij}. With all evidence (i, j) pairs ranked for $i = 1, ... , I/2$ and $j = 1, ... , J$, sum the corresponding arrival proportions, starting with the highest-ranking (i, j) pair and adding the next highest-ranking (i, j) pair one at a time, until the sum equals or exceeds c/a, which is the proportion of arrivals that can be processed. All (i, j) pairs in this sum are processed (although the last (i, j) pair included in the sum may be only partially processed), and all (i, j) pairs not in this sum are not processed.

Case 2: The arrival rate of evidence, a_e, is less than the capacity, c. In this case, all evidence is processed along with a portion (the ones with the highest hit proportion) of test-fires. Rank all test-fire (i, j) pairs for $i = I/2 + 1, ... , I$ and $j = 1, ... , J$ by the hit probability h_{ij}. Sum their corresponding arrival proportions, starting with the highest-ranking (i, j) pair and adding the next highest-ranking (i, j) pair one at a time, until the sum equals or exceeds $(c - a_e)/a$. All (i, j) test-fire pairs in this sum are processed, and all (i, j) test-fire pairs not in this sum are not processed.

Comment 1: If it is desired to give homicides strict priority over non-homicides in the capacity allocation process, then only include non-homicide cartridge cases in this procedure, and let c be the available capacity for nonhomicide cartridge cases.

Comment 2: If the data required to estimate the hit probabilities h_{ij} are unavailable, then replace h_j in Step 4 with $\sum_{i=1}^{I} a_{ij}$, which is the proportion of arrivals that are of caliber

j, and assume all (*i, j*) pairs have the same value of h_{ij} in Step 5. This essentially assumes that more popular calibers have higher hit probabilities. In WBW, the correlation between $(h_1, \ldots, h_{12})$ and $(\sum_{i=1}^{I} a_{i,1}, \ldots, \sum_{i=1}^{I} a_{i,12})$ is 0.56, which justifies this substitution.

Conclusion

Data from Stockton, CA, which processes all of its cartridge cases for ballistic imaging purposes, provide an uncensored view that allows for the counterfactual estimation of what would happen if processing capacity was limited. This study, which summarizes work in WBW, has two main results. First, when capacity is limited and the Random Policy is used, the number of hits increases according to *(capacity ÷ arrival rate)²* , which is much smaller than in traditional service and manufacturing operations, where output increases according to *(capacity ÷ arrival rate)*. The value of *(capacity ÷ arrival rate)* varies widely among different U.S. agencies [5]. If poorly performing sites (i.e., those that process few cartridge cases and generate few hits [5]) linearly extrapolate what their performance would be if they increased their processing capacity, they would be grossly underestimating the efficacy of ballistic imaging systems. Put another way, this derivation that ballistic imaging has increasing returns to scale (as exhibited in the convexity of the lowest curve in **Figure 1**) should motivate poorly performing sites to invest more in ballistic imaging.

The second main result is the proposal of a simple capacity allocation policy that combines three ideas: prioritizing evidence over test-fires, grouping evidence cartridge cases by caliber to eliminate the convexity plaguing the lowest curve in **Figure 1**, and prioritizing entry type/caliber pairs that have high hit probabilities. This policy is easy and inexpensive to implement (and performs well even if the hit probabilities for entry type/ caliber pairs are unknown), and can significantly increase the number of hits; e.g., the hit rate can be approximately doubled when the processing capacity is half of the arrival rate.

As with any mathematical model, there are some limitations. First, aside from the possibility of prioritizing homicides over non-homicides, the model does not consider the "usefulness" of a hit, beyond the fact that it occurred. See WBW for a comparison of using Last-Come First-Served vs. First-Come First-Served to maximize the number of hits that are obtained before the corresponding criminal case is closed. Another limitation is that the model assumes that test-fires require no additional processing. Incorporating the possibility that agencies perform their own test-firing would not alter the proposed capacity allocation policy because test-fires are already awarded lower priority than crime scene evidence. Finally, the model assumes that the rate of crime and the mix of entry type/caliber do not vary over time. Parameter re-estimation should be performed periodically, especially if there is a change in policing strategy. Parameter re-estimation may be particularly important for entry sites that perform entries for multiple outside agencies because they may not necessarily receive information about the need for a reassessment from these outside agencies.

Acknowledgments

References

[1] Wang, C., Beggs-Cassin, M., and Wein, L.M., "Optimizing Ballistic Imaging Operations," Journal of Forensic Sciences, Vol. 62, No. 5, Sept. 2017, pp. 1188-1196.

[2] Cork, D.L., Rolph, J.E., Meieran, E.S., and Petrie, C.V., editors, "Ballistic Imaging," National Academies Press, Washington, DC, 2008.

[3] Office of the Inspector General, U.S. Department of Justice, "The Bureau of Alcohol, Tobacco, Firearms and Explosives' National Integrated Ballistic Information Network Program," Audit Report 05-30, Office of the Inspector General, U.S. Department of Justice, Washington, DC, 2005.

[4] King, W., Wells, W., Katz, C., Maguire, E., and Frank, J., "Opening the Black Box of NIBIN: A Descriptive Process and Outcome Evaluation of the Use of NIBIN and its Effects on Criminal Investigations," U.S. Department of Justice, National Institute of Justice, Washington, DC, 2013.

[5] Maguire, E.R., King, W.R., Matasiuak, M.C., and Campbell, B., "Testing the Effects of New Personnel, Processes, and Technology on Ballistic Evidence Processing Productivity," Police Quarterly, Vol. 19, 2016, pp. 199-215.