2013

An approach for criminal career analysis using hazard patterns

Carl A. Janzen
Dakota State University

Amit Deokar
Dakota State University

Omar F. El-Gayar
Dakota State University

An approach for criminal career analysis using hazard patterns

Conference Paper · December 2013

CITATIONS
0

READS
76

3 authors:

Carl Janzen
University of the Fraser Valley
3 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Amit Deokar
University of Massachusetts Lowell
89 PUBLICATIONS **556** CITATIONS
SEE PROFILE

Omar F. El-Gayar
Dakota State University
156 PUBLICATIONS **1,500** CITATIONS
SEE PROFILE

An approach for criminal career analysis using hazard patterns

1.Introduction

The facilities in California's prison system were designed to house approximately 85,000 inmates. These facilities held approximately 156,000 inmates in 2011, when the Supreme Court upheld an order that would require the state to decrease the prison population by 46,000 (Newman & Scott, 2012). The court mandated the release of tens of thousands of inmates, because the prison system was unable to provide adequate health care to the inmate population (Bower, 2012).

One key way to reduce the number of individuals serving their sentence in prison is through parole release. However, identifying candidates for successful parole release is no easy task when recidivism rate is high and the number of lifelong desisters is low. In the context of a criminal career, recidivism is the re-occurrence of an arrest charge or conviction, while desistance is the absence of such a re-occurrence. The rate of recidivism will vary depending on whether the subject of interest is arrest or conviction. In California, 84% of individuals released from prison during the fiscal year 2007-2008 were re-arrested within three years of release, and 60% were convicted (Cate et al., 2012).

There are two primary goals for this work. The first goal is to find how we can assess risk of recidivism based on past offending behavior. The second goal is to codify the risk, as well as the basis for the risk in simple terms.

In this paper, we make two key contributions. First, we demonstrate that hazard patterns can be used to discover patterns that can reliably predict differences in risk of re-arrest following parole release. Second, we propose and demonstrate a test of meaningfulness for hazard patterns. Without such a test, it can be difficult to differentiate between patterns that occur by chance and genuine meaningful patterns.

The remainder of the paper is organized as follows. In section 2, we present the problem in the context of related work. In section 3, we describe the needs that a solution to the presenting problem should address. In section 4 we describe the design of the system through use of an illustrative example. In section 5 we demonstrate and evaluate the pattern discovery system, and finally in section 6 we summarize our evaluation, describe limitations, discuss implications for predicting recidivism, and identify directions for future work.

2.Related Work

In this section, we discuss the problem context and the motivation for this work. We provide a review of relevant criminology literature with attention to recidivism prediction in parolees.

2.1 Risk Assessment

A number of state-specific risk assessment systems have been developed to address this need. Examples include Ohio's progressive sanction grid (Martin & Dine, 2008), the Minnesota Screening Tool Assessing Recidivism Risk (MnSTARR) (Duwe, 2013), and the California Parole Violation Decision Making Instrument (PVDMI) (Turner, Braithwaite, Kearney, Murphy, & Haerle, 2012).

Both the Ohio screening tool and the PVDMI encountered considerable resistance from practitioners in the field. Parole officers using the Ohio screening tool questioned whether the decisions of the tool were logical, and parole officers in California consistently escalated the recommended sanction for parolees with significant prior criminal behavior. Turner et al. (2012) suggested that parole officers may not have been confident that criminal histories were properly taken into account by the system. It remains to be seen whether the deployment of MnSTARR will fare better. Resistance to actuarial risk assessment tools is not altogether unjustified. A meta-analysis 1 of risk assessment instruments found they produced an area under curve (AUC) scores rang- ing from 0.65 to 0.71(Min, Wong, & Coid, 2010).

2.2 Criminal career analysis

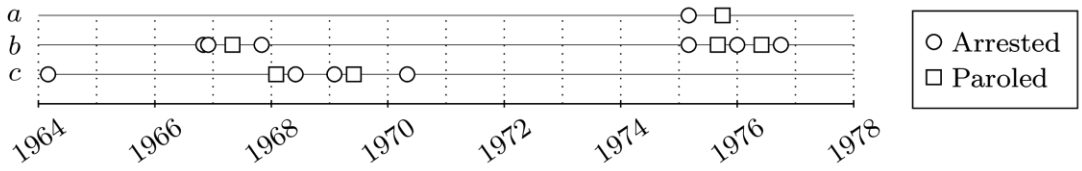

Figure 1: contrived histories

Research in the area of quantitative criminal career analysis deals specifically with criminal histories, and commonly makes use of group trajectory modeling. This technique was first introduced by Nagin and Land (1993) and has since been used in many other studies. Bhati and Piquero (2007) supplemented this technique with variables representing the time between preceding arrest incidents, to more accurately predict future recidivism. However, predicting future arrests based on group trajectories is still a difficult task. Researchers have cautioned against policy development based on such tools, and have suggested that improvements will not be obtained using new analysis methods (Bersani, Nieuwbeerta, & Laub, 2009).

3.Objectives of a Solution

Based on the process review of the PVDMI pilot deployment in California, practitioners lacked confidence in the logic supporting the tool's risk determinations, and did not believe the tool properly accounted for changes in risk associated with repeat offending behavior (Turner et al., 2012). We are presented with the challenge, not only of assessing risk, but of justifying that risk assessment to a decision maker, particularly with respect to prior offending record. We have identified two key design requirements for a recidivism risk assessment tool:

1. Incorporate salient characteristics of prior record to determine risk.
2. Concisely present the logic leading to the determined risk level.

One way to represent a criminal history is as an ordered sequence of many different types of events. Event sequences that occur frequently can be represented as patterns for classification, clustering, or prediction tasks. Hazard patterns are frequent sequences of events where each successive event in a pattern represents the first subsequent event of that type, and where the time between events in a pattern represents time-to-failure or time-to- event (Janzen, Deokar, & El-Gayar, 2013b). As already noted in (Bhati & Piquero, 2007), time between preceding arrests is a useful predictor of future arrest risk. A hazard pattern representing a history of many arrest charges for various offenses will also capture periods of desistance, during which no arrest occurred. Hazard patterns draw on survival analysis techniques, and allow the analyst to include potentially significant in- formation about time between events. However, to demonstrate usefulness and reliability of these patterns for risk assessment, we must address some important concerns:

1. **Over-fitting:** Are the patterns generalizable to other similar data sets?
2. **Meaningfulness:** Are patterns found even when there are no patterns in the data?
3. **Predictiveness:** Can historical patterns be useful predictors of future events?
4. **Parsimony:** Can we produce output with minimal redundancy?

A common way to address the concern of over-fitting is to rely on some form of validation on a hold-out sample. One portion of the data set is set aside only for validation, while the remainder of the dataset is used to train the model. Repeating this process k times, the data is divided in to k subsets, each of which serves as a validation set for a model trained using the remainder of the data.

The second issue of meaningfulness is both subtle and important. Keogh and Lin (2004) presented the surprising result that a subsequence clustering technique used in dozens of published papers produced meaningless results. For our test, we draw on their formal definition: "We call an algorithm

meaningless if the output is independent of the input." We will test for meaningless results by altering salient characteristics of the input, to provide evidence to support the belief that the discovered patterns are meaningful.

The third concern of predictiveness is of vital importance. We can demonstrate predictiveness by showing that patterns discovered in one time period can reliably predict arrest risk in a subsequent time period. This is also the most difficult test, since patterns learned in the past cannot account for future changes in the environment.

Finally, to avoid producing an overwhelming number of patterns that may or may not be useful, we must apply a pattern selection strategy.

4.Design and Development

In this section we describe the system design by example, using a collection of five contrived criminal histories. We first present the events on a time line, and then refer to this example as we describe our search for a solution.

4.1 Hazard Patterns

Figure 1 contains three contrived example criminal histories for individuals a, b and c. This example is simplified for the sake of illustration. Individuals b and c are arrested and re-released on parole more than once, while a is released on parole only once and is not re-arrested. Keeping in mind that these are contrived histories, can we find a relationship between patterns of past event occurrences and risk of arrest after parole release?

An **event occurrence** is denoted (e, t) where e represents the event type and t represents the time of the event occurrence. For example, $(Paroled, 909)$ is the occurrence of event $Paroled$ at time 909 (months since January 1900).

An **event sequence** of length n is denoted $\langle (e_1, t_1), (e_2, t_2), \ldots (e_n, t_n) \rangle$ where e_i rep- resents the type of the i^{th} event, t_i represents the time of the i^{th} event, and $t_{i-1} < t_i$. An event sequence is a time oriented arrangement of event occurrences. For example, $\langle (Paroled, 833), (Arrested, 844) \rangle$ is an event sequence.

Event sequence mining has been used to discover characteristic patterns for classification (Zaki, Lesh, & Ogihara, 1998), clustering (Bathoorn, Welten, & Richardson, 2010), and pattern discovery (Fujikawa, Kida, & Katoh, 2011).

As event histories increase in length, the probability of discovering frequent event sequences by chance increases. To counter this effect, it is useful to apply a gap constraint on the time between events of interest.

A **gap constraint** is the requirement that except for the initial event occurrence, for any event occurrence (e_i, t_i) there exists at least an occurrence (e_{i-1}, t_{i-1}) where $mingap \leq (t_i - t_{i-1}) \leq maxgap$. Two events in an event sequence satisfy a minimum gap constraint if they are separated by at least $mingap$ and they satisfy a maximum gap constraint if they are separated by at most $maxgap$. For a more detailed discussion see (Leleu, Rigotti, Boulicaut, & Euvrard, 2003).

However, gap constraints do not include information about whether the event of interest occurs additional times prior to $mingap$. To describe a relationship between previous events and time to re-arrest, hazard patterns and hazard constraints were proposed in (Janzen et al., 2013b).

A **hazard pattern** is a frequently occurring sequence of events where each subsequent event occurrence is the first subsequent occurrence of that particular event type. A hazard pattern can be denoted as $Paroled \rightarrow Arrested$. For all occurrences of this pattern, $Arrested$ refers to the first arrest after parole release.

We can further apply a **hazard constraint** whereby the period of time between two events in a pattern must fall within a specified minimum and maximum time interval. A given hazard pattern

Table 1: Months until re-arrest (contrived data set)

Antecedents	$\overrightarrow{(0,3]}$	$\overrightarrow{(3,6]}$	$\overrightarrow{(6,12]}$	$\overrightarrow{(12,24]}$	$\overrightarrow{(24,96]}$	$\overrightarrow{(96,384]}$	Total
		RR 1.50					
		Z 0.53					
Arrested $\overrightarrow{(3,6]}$ *Paroled*	0.00	0.75	0.25	0.00	0.00	0.00	1.00
support 5		s 3	s 1				s 4
distinct 4		d 3	d 1				d 4
individuals 2		i 1	i 1				i 2

$Paroled \rightarrow Arrested$ can be expressed with a hazard constraint as $Paroled\ \overrightarrow{(minhaz, maxhaz]}\ Arrested$. Occurrences of $Paroled\ \overrightarrow{(3,6]}\ Arrested$ satisfy the condition that more than three and at most six months elapsed between parole release and the next arrest.

A straightforward way to describe relationships between antecedent patterns and subsequent events is to describe the proportion of antecedents that lead to the consequent. There are six distinct occurrences of Paroled. Of these six occurrences, we see in table 1 that four lead to re-arrest within 4-6 months, a proportion of 0.67 re-arrests per parole release for that time period. Counting the number of occurrences is less straightforward when more than one antecedent event leads to the same subsequent. In table 1 we see that Arrested → Paroled occurs four distinct times but has a support count of five. This is because more than one Arrested event leads to the same subsequent Paroled event. In event sequence mining there is no agreed upon way to count the number of pattern occurrences. For instance, (Achar, Laxman, & Sastry, 2011) describes 10 different support counting methods.

For all support counting methods we encountered, one or more of the following were true: (a) counts were non-independent of other occurrences of the same pattern (non-overlapping, non-interleaved, and distinct occurrence based), (b) longer patterns were unduly penalized (window and expiry time based) and (c) unrelated event occurrences can inflate support counts (head frequency, total frequency). For further discussion of these limitations, see (Janzen et al., 2013a).

To be able to adequately express the relationship between the antecedent and the sub- sequent, rather than counting the number of times the antecedent and subsequent occur together, we counted the number of antecedents that lead to the subsequent.

Relative Support is the number of distinct or unique antecedent event occurrences that are followed by a subsequent event of a particular type in a hazard pattern. In table 1, the an- tecedent pattern $Arrested\ \overrightarrow{(3,6]}\ Paroled$ has a support of 5, but we only consider 4 distinct an- tecedents when calculating the proportion that participates in $Arrested\ \overrightarrow{(3,6]}\ Paroled\ \overrightarrow{(3,6]}\ Arrested$. Relative support was proposed for event hazard patterns in (Janzen et al., 2013b).

An additional problem is the large number of patterns discovered. To determine whether a particular pattern might convey useful information, we can calculate a measure of interest and apply a statistical test of significance.

Relative Risk is the ratio of the risk within a treatment group over the risk of the control group. It is used to measure the cumulative treatment effect at the end of a period of time. For a practical discussion of relative risk ratios, see (Bewick, Cheek, & Ball, 2004).

We evaluated pattern selection using significance tests on Relative Risk (RR). Patterns shown to significantly affect the RR coefficient in training data were also shown to have a similar effect in test data. For further details, see (Janzen, Deokar, & El-Gayar, 2013a).

RR expresses the ratio between survival proportions in a treatment group compared to the same in a control group. We compare the RR for a presented pattern with the RR for the same pattern with the first antecedent removed. In table 1, RR could only be calculated in this way for one of the patterns. The risk of arrest in the four distinct antecedent parole releases in

encoded	event
0	Arrested
1	Paroled

(a) events

encoded	constraint
0	(0, 0]
1	(0, 3]
2	(3, 6]
3	(6, 12]
4	(12, 24]
5	(24, 96]
6	(96, 384]

(b) constraints

Ordinal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Arrested	0	0	3	5	5	6	8	8	10	10	0	13	13	14	16	16	0
Paroled	1	0	4	4	7	7	7	9	9	0	0	12	15	15	15	0	0

(c) ordinal index

Ordinal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Arrested	0	0	1	3	2	5	3	2	3	2	0	5	2	3	4	3	0
Paroled	3	0	2	2	6	5	2	3	2	0	0	5	4	3	2	0	0

(d) constraint index

ordinal	offset	individual
0	902	0
1	909	0
2	802	1
3	803	1
4	808	1
5	814	1
6	902	1
7	908	1
8	912	1
9	917	1
10	921	1
11	770	2
12	817	2
13	821	2
14	829	2
15	833	2
16	844	2

(e) offsets

Figure 2: data structures

$Arrested\ \overrightarrow{(3,6]}\ Paroled$ (3/4) against the risk of arrest in the two distinct parole releases in Paroled that are not already counted in $Arrested\ \overrightarrow{(3,6]}\ Paroled$ (1/2). The RR of 1.50 indicates that the risk of re-arrest during the subsequent 4-6 months is one and a half times higher if parole release is 4-6 months after arrest. A RR of 1 would indicate no change. To see whether the increase in risk might be generalizable to the broader population, we calculate a Z-score for the RR. In this case, the resulting Z-score of 0.53 indicates no evidence to expect that RR is different than 1.

5. Algorithm Design

5.1.1 Data Structures

To facilitate indexing, constraint and offset values were stored in a lookup table. Events were encoded as integers, constraints of increasing sizes were represented as successive integers, and offset values were represented as ordinals. Offsets were kept separated per individual, as illustrated in figure 2(e)

Using these simplified representations of events, offsets and constraints, an index was constructed to enable easy lookup of both when the next event of a given type might occur, as well as what constraint is satisfied by that occurrence.

For instance, by referencing ordinal 6 in figure 2(c) we see in the ordinal index that the next Arrested event occurs at ordinal 8 and we see in the constraint index that constraint 3 is satisfied for that next ordinal. A value of zero indicates that there is no applicable next ordinal. Note that histories of different individuals are indexed back to back, index columns 1,10, and 16 contain zeroes because the subsequent ordinal belongs to a different individual's history.

We can also see convergence from multiple antecedents to a single subsequent. For columns 5 and 6, we first see that two distinct Arrested events occurred. For each of the arrest events, the next Arrested event is a different occurrence. However, in the bottom row, we see that for both of these, the next Paroled event is the same one, at ordinal 7. This convergence is also seen in figure 1 in individual b at the end of 1966. The algorithms used to construct the above indexes and lookup tables are not detailed here.

5.1.2 Pattern discovery algorithm

The discovery of frequent patterns can follow either a depth first or breadth-first tree traversal, due to lack of dependencies between branches. Frequent antecedent ordinals are collected, and for each type of subsequent event, the ordinals are grouped according to the constraint they satisfy (each ordinal satisfies only one constraint). Constraint groupings that are larger than a specified support threshold become candidates for further extension. The

ALGORITHM 1: Grow()

required:
ordinal database D
constraint index G
$\forall(event, ordinal) \in D : event \in M$
ensure : $\forall pat \in Freq : |pat| \geq sup$
input : $pat, Ords$

```
1  for event ∈ M do
2      for constrt ∈ G do
3          if |Ords_{constrt,event}| ≥ sup then
4              newpat ← (pat, constrt, event)
5              append newpat to Freq
6              NextOrds ← Next(Ords_{constrt,event})
7              Grow(newpat, NextOrds)
8          end
9      end
10 end
```

ALGORITHM 2: Next()

required:
$[R_{event,ordinal}]_{(m\times n)}, [I_{event,ordinal}]_{(m\times n)}$
alphabet of all event types M
input : $Antecedents$
output : $Subsequents$

```
1  Antecedents ← Unique(Antecedents)
2  for event ∈ M do
3      for ord ∈ Antecedents do
4          c ← R_{event,ord}  nextOrd ← I_{event,ord}
5          append nextOrd to Subsequents_{event,c}
6      end
7  end
8  return Subsequents
```

Grow function shown in algorithm 1 relies on the above indexes. Ordinals are translated to offsets at $O(1)$ cost as needed. Input ordinals are supplied in a matrix indexed by event, constraint, where each $M_{event,constraint}$ represents the antecedent ordinals for the current pattern growth step. In Line 3, those antecedents with cardinality that is high enough to meet a specified support threshold are added to the frequent pattern database in line 5, and are passed to the $Next$ function, where a new matrix of candidate ordinals is created, and passed to the recursive $Grow$ attempt on line 7.

Algorithm 2 contains the $Next$ function, which takes as input a collection of antecedent ordinals, grouped by event, and produces the Ordinal matrix $NextOrds$ needed in line 6 of algorithm 1. This function uses two indexes: $R_{event,ordinal}$ and $I_{event,ordinal}$. Refer to the indexes and lookup tables in figure 4.2.1. R and I are matrices of dimension $(m \times n)$ where m is the alphabet of all possible events, and n is the number of distinct offsets. Multiple events may occur at the same offset. I contains the ordinal of the subsequent occurrence of a given event type. R contains the constraint that is satisfied at a given event offset (represented as an ordinal), relative to the immediate antecedent event.

On line 4 of the $Next$ function pseudo-code in algorithm 2, for each antecedent event occurrence, the constraint $R_{event,ordinal}$ that is satisfied for each potential subsequent event is retrieved. Given the half-open interval topology used to describe the different constraints, each subsequent event can satisfy one constraint. In line 4 the subsequent ordinals are retrieved from I and then grouped according to their matching constraints in line 5. The creation of R and I are not described here, but are straightforward.

6.Evaluation

In this section, we demonstrate the results obtained using the pattern discovery system to mine a data set of real life criminal histories. We then evaluate the pattern discovery system according to the four design objectives described in section 3: over-fitting, meaningfulness, predictiveness, and parsimony.

The pattern discovery system was used to discover patterns in a data set of complete criminal histories. The histories were collected from part of a non-random sample of offenders who entered the California Youth Authority's Deuel Vocational Institute in 1964 and 1965. The event database contains 54,175 arrest records and associated dispositions, parole, and discharge events for 3,652 individuals from the time of first arrest through 1983. Dates were discretized to the nearest 15th day of the month (Wenk, 2006).

All dispositions (including convictions) were recoded to the arrest charge date. Any patterns showing both arrests and convictions have nothing to do with conviction rates.

We selected patterns with a minimum support of 500, as well as related stub patterns. Stub patterns are those patterns that would otherwise be excluded due to low support, but which are

Table 2: Arrests after parole

Antecedents	$\overrightarrow{(0,3]}$	$\overrightarrow{(3,6]}$	$\overrightarrow{(6,12]}$	$\overrightarrow{(12,24]}$	$\overrightarrow{(24,96]}$	$\overrightarrow{(96,384]}$	Total
	RR 1.13	RR 1.04	RR 1.01	RR 1.02	RR 0.95	RR 0.85	
	Z **3.54**	Z 0.97	Z 0.28	Z 0.40	Z -0.71	Z -0.68	
Convicted $\overrightarrow{(6,12]}$*Paroled*	0.30	0.21	0.20	0.15	0.10	0.01	0.96
support 4115	s 1053	s 717	s 704	s 528	s 334	s 29	s 3365
distinct 3497	d 1053	d 717	d 704	d 528	d 334	d 29	d 3365
individuals 2364	i 863	i 643	i 645	i 496	i 330	i 29	i 2261
	RR 1.11	RR 1.05	RR 1.00	RR 0.79	RR 0.89	RR 1.61	
	Z **3.06**	Z 1.16	Z -0.07	Z **-4.45**	Z -1.87	Z **2.13**	
Convicted $\overrightarrow{(12,24]}$*Paroled*	0.30	0.21	0.20	0.13	0.09	0.01	0.94
support 5885	s 1210	s 839	s 815	s 536	s 376	s 47	s 3823
distinct 4087	d 1210	d 839	d 815	d 536	d 376	d 47	d 3823
individuals 2720	i 981	i 755	i 739	i 508	i 368	i 47	i 2563
	RR 0.87	RR 0.97	RR 0.93	RR 0.99	RR 1.18	RR 0.73	
	Z **-3.88**	Z -0.74	Z -1.61	Z -0.19	Z **2.46**	Z -1.28	
Convicted $\overrightarrow{(24,96]}$*Paroled*	0.25	0.20	0.19	0.15	0.11	0.01	0.91
support 6871	s 774	s 595	s 579	s 451	s 332	s 23	s 2754
distinct 3042	d 774	d 595	d 579	d 451	d 332	d 23	d 2754
individuals 2038	i 642	i 529	i 528	i 428	i 328	i 23	i 1885
	RR 1.48	RR 1.18	RR 0.97	RR 0.62	RR 0.58	RR 0.45	
	Z **10.28**	Z **3.17**	Z -0.59	Z **-5.81**	Z **-5.16**	Z **-2.03**	
Paroled $\overrightarrow{(12,24]}$*Paroled*	0.38	0.23	0.19	0.10	0.06	0.00	0.97
support 1504	s 577	s 345	s 292	s 149	s 92	s 7	s 1462
distinct 1504	d 577	d 345	d 292	d 149	d 92	d 7	d 1462
individuals 1065	i 473	i 306	i 269	i 146	i 91	i 7	i 1042
	RR 1.04	RR 1.02	RR 0.93	RR 1.00	RR 0.89	RR 0.89	
	Z 1.22	Z 0.39	Z -1.57	Z -0.06	Z -1.59	Z -0.48	
Paroled $\overrightarrow{(24,96]}$*Paroled*	0.29	0.20	0.19	0.15	0.09	0.01	0.93
support 2886	s 833	s 584	s 549	s 430	s 263	s 25	s 2684
distinct 2886	d 833	d 584	d 549	d 430	d 263	d 24	d 2683
individuals 1682	i 655	i 489	i 478	i 388	i 258	i 24	i 1592

siblings of a frequent pattern. For instance, if the subsequent event occurs frequently in the follow-up period of (0,3], we also tabulate the number of occurrences in the adjacent follow-up periods, and calculate a total. Table 2 shows only patterns with an antecedent ending with parole and a subsequent event of Arrest. We see that the recidivism is generally high in this group.

We note several relationships between criminal history and recidivism. Of all the follow-up periods, even though the (0,3] time interval is the smallest, it also tends to be the time period with the highest support counts. Over all, there is only a small amount of variation between the groups represented by each pattern. RR values for shorter follow-up periods are closer to 1, with larger Z-scores, and RR values for longer follow-up periods are farther from 1, with smaller Z-scores. The patterns provide more generalizable information about the short follow-up periods. Past repeat offending over (0,3] increases risk of repeating the same when released on parole after (12,24]. Generally, those who are released on parole sooner also re-offend sooner than others. The single strongest relationship is shown in the second-last row. Time since previous parole release has a large and significant impact on recidivism. Individuals released on parole (12,24] after their previous parole release are 1.48 times as likely to re-offend within 3 months when compared to all others released on parole. In the last row, we see no significant change in RR for antecedent parole releases (24,96] apart. Using the same data set, mined at a lower minimum support threshold, we observed other patterns relating to specific arrest charges, dismissals, and convictions.

6.1 Evaluation against design goals

Over-fitting: In the case of a very complex pattern discovery system, it may be possible to over-fit the characteristics of the training set. To test against this, we performed a *k*-fold cross-validation with ten folds. Each fold consisted of a 90% training split and a 10% testing split. We selected patterns based on a RR Z score outside ±1.96. For each fold, we considered contradictions to be those cases where the training split and the testing split each reported a significant Z score of opposite sign. We recorded consistency where a significant Z score in the testing split corresponded to a Z score of the same sign in the training split. For each fold, pattern mining was performed with a

Shuffled	Fold	Patterns	Sig	Contradict	Consistent
False	0	11756	3551	0.00	0.97
False	1	12163	3510	0.00	0.97
False	2	11633	3766	0.00	0.98
False	3	11788	3649	0.00	0.97
False	4	11420	3682	0.00	0.97
False	5	11770	3690	0.00	0.97
False	6	11925	3652	0.01	0.97
False	7	12031	3549	0.01	0.97
False	8	11837	3549	0.00	0.97
False	9	11508	3503	0.00	0.97
True	0	5903	1362	0.01	0.96
True	1	5919	1493	0.01	0.94
True	2	5872	1556	0.01	0.95
True	3	5877	1552	0.01	0.95
True	4	5741	1351	0.01	0.95
True	5	5880	1521	0.01	0.95
True	6	5980	1446	0.01	0.96
True	7	5995	1347	0.01	0.96
True	8	5860	1500	0.01	0.95
True	9	5752	1428	0.01	0.95

(a) Cross Validation

Year	prop Train	prop Test	Z Prop	RR Train	RR Test	Z RR
61	0.42	0.96	**6.20**	**0.67**	**1.06**	**2.21**
62	0.69	0.98	**8.24**	**0.98**	**1.03**	**2.70**
63	0.87	0.94	**4.23**	**1.16**	**1.09**	**1.03**
64	0.90	0.90	0.40	1.12	1.12	0.09
65	0.94	0.89	**-7.82**	**1.08**	**1.11**	**-1.52**
66	0.97	0.90	**-12.16**	**1.06**	**1.10**	**-2.24**
67	0.94	0.92	**-3.39**	**1.08**	**1.07**	**-0.66**
68	0.92	0.93	0.35	1.09	1.10	0.07
69	0.91	0.92	0.49	1.10	1.13	0.10
70	0.90	0.92	1.67	1.11	1.15	0.35
71	0.90	0.89	-0.96	1.11	1.11	-0.23
72	0.90	0.85	**-3.45**	**1.10**	**1.11**	**-0.92**
73	0.92	0.85	**-3.93**	**1.10**	**1.10**	**-1.03**
74	0.92	0.86	**-3.10**	**1.12**	**1.06**	**-0.79**
75	0.90	0.91	0.52	1.12	1.14	0.11
76	0.89	0.95	**4.53**	**1.12**	**1.26**	**0.86**
77	0.87	0.94	**4.34**	**1.11**	**1.29**	**0.90**
78	0.86	0.87	0.61	1.08	1.19	0.16
79	0.89	0.79	**-2.88**	**1.12**	**1.12**	**-0.87**
80	0.89	0.71	**-4.25**	**1.12**	**1.21**	**-1.43**

(b) Prediction

Figure 3: evaluation results

Table 3: Arrests after parole - patterns found in shuffled data

Antecedents	$\overrightarrow{(0,3]}$	$\overrightarrow{(3,6]}$	$\overrightarrow{(6,12]}$	$\overrightarrow{(12,24]}$	$\overrightarrow{(24,96]}$	$\overrightarrow{(96,384]}$	Total
	RR 0.98 Z -0.50	RR 0.98 Z -0.41	RR 0.89 Z **-2.30**	RR 0.82 Z **-4.41**	RR 1.13 Z **2.45**	RR 1.57 Z **2.03**	
Convicted $\overrightarrow{(24,96]}$ *Paroled*	0.19	0.13	0.16	0.18	0.18	0.01	0.85
support 6657	s 604	s 433	s 523	s 575	s 577	s 40	s 2752
distinct 3219	d 604	d 433	d 523	d 575	d 575	d 40	d 2750
individuals 2165	i 557	i 402	i 489	i 543	i 551	i 40	i 1921

minimum support threshold of 500. We tabulated the above indicators for each of the ten folds, and Type equation here.repeated the process for the same 500 individuals with all events shuffled.

Since each training split was much larger than each corresponding test split we compared the sign of the significant Z-scores (±1.96) for patterns in the test split with the sign of the corresponding patterns in the training split. A shown in figure 3(a), we observed good consistency and few contradictions.

Meaningfulness: Our next concern was whether the discovered patterns were meaningful. More generally, how will we know whether the discovered patterns are simply an artifact of the mining process? If we discover more hazard patterns in shuffled data than can be expected by random chance alone, then the minimum support threshold is too low. For the purpose of the presenting problem, patterns can be considered meaningful if the number of patterns found in a shuffled equivalent of the data are markedly different than those found in the original data. When we mined for patterns in randomly shuffled data, we still discovered hazard patterns, but in much smaller numbers, as shown in figure 3(a).

We also repeated the pattern mining and pattern selection process used for table 2, using the same data, but shuffled. Rather than 10 significant patterns discovered, there were only 4 significant patterns discovered in the shuffled data set (see table 3 allowing us to conclude that the patterns shown in table 2 are indeed meaningful.

Predictiveness: To evaluate the predictiveness of the discovered patterns, we first compared the proportion of patterns that lead to arrest in one time period with the proportion of pattern occurrences that lead to arrest in a subsequent time period. We evaluated the entire time period for the antecedent pattern $Paroled\ \overrightarrow{(12,24]}\ Paroled$ to predict re-arrest within a two 8 year time span. If there is no significant difference between the arrest risk over a preceding time period and a subsequent time

period, then past risk of offending, the hazard pattern, might be a useful predictor of future risk. We also computed RR for each of the time periods, comparing the arrest risk for $Paroled\ \overrightarrow{(12,24]}\ Paroled$ with a baseline arrest risk for $Paroled$. For each parole release, we tabulated the number of re-arrests within 2 years of those parolees released 24- 48 months prior (training period). We compared that proportion with the proportion of re-arrests within two years going forward (testing period). To reduce variance, and to summarize the results, these were grouped according to year, as shown in figure 3(b). Some arrests before 63/64 correspond to juvenile offenses. This corresponds to the most difficult time period for predicting arrest risk for these individuals. We anticipate that prediction accuracy would further improve with a data set comprised of individuals with varying ages.

We noted that past risk of arrest is significantly different than future risk of arrest in 13 of 20 years. However, past relative risk of arrest was not significantly different than future relative risk, except in 3 of 20 years ($Z \pm 1.96$). We repeated the test using a variety of different training, testing and follow-up periods. In each case, the results were similar. For this particular hazard pattern, we observed that past risk of arrest is not a reliable indicator of future risk of arrest. Further, past relative risk is a good indicator of future relative risk. In other words, the RR between $Paroled\ \overrightarrow{(12,24]}\ Paroled$ and $Paroled$ does not significantly change over the selected time periods. Based on these observations, RR of hazard patterns is a useful measure for identifying parolees who are at relatively higher risk to re-offend.

Parsimony: In table 2 we summarized all patterns related to recidivism after parole release with a minimum support threshold of 500, and logically arranged them together with indicators of effect direction and significance in bold. Based on the above pattern tables, we see that the use of the RR Z-score dramatically reduces the number of patterns of interest, reducing the number of patterns for the analyst to consider.

7.Conclusion

In this paper we demonstrated that hazard patterns can be used to identify individuals with increased parole violation risk. Although we did not find a direct link between past arrest risk and future arrest risk, we did find a significant relationship between past RR and future RR between a group exhibiting a hazard pattern and a group that exhibited a hazard pattern with the first antecedent removed. We also tested the generalizability of the discovered hazard patterns through ten-fold cross validation. We further also demonstrated a simple test for meaningfulness of hazard patterns. If a similar amount of patterns is discovered when the order of the underlying data is shuffled, then the discovered patterns are meaningless. The need for a meaningfulness test is particularly relevant, given that meaningless patterns can pass a cross-validation test. An important limitation is a result of multiple testing. Since many tests of significance were performed, an analyst may encounter one of many possible relationships by chance alone. We did not examine the impact of multiple testing bias in this work. With the introduction of hazard patterns comes a wide range of opportunities for further work. Application domains with time-to-event data are the most likely to benefit. Examples include health care histories, business process analytics, equipment failure events, and insurance claim histories.

References

Achar, A., Laxman, S., & Sastry, P. S. (2011). A unified view of the apriori-based algorithms for frequent episode discovery. *Knowledge and Information Systems*. doi: 10.1007/s10115-011-0408-2

Bathoorn, R., Welten, M., & Richardson, M. (2010). Frequent episode mining to support pattern analysis in developmental biology. *Pattern Recognition in bioinformatics*.

Bersani, B. E., Nieuwbeerta, P., & Laub, J. H. (2009). Predicting Trajectories of Offending over the Life Course: Findings from a Dutch Conviction Cohort. *Journal of Research in Crime and Delinquency, 46*, 468-494. doi: 10.1177/0022427809341939

Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 11: assessing risk. *Critical care (London, England), 8*, 287-291. doi: 10.1186/cc2908

Bhati, A. S., & Piquero, A. R. (2007). Estimating the Impact of Incarceration on Subsequent Offending Trajectories: Deterrent, Criminogenic, or Null Effect? *The Journal of Criminal Law and Criminology, 98*, 207-253.

Bower, A. (2012). Unconstitutionally crowded: Brown v. Plata and how the Supreme Court pushed back to keep prison reform litigation alive. *Loy. LAL Rev., 45*.

Cate, M. L., Hoshini, M., Seale, L., Grealish, B., Fitzgerald, T., Grassel, K., . . . Reyes, M. (2012). *2012 CDCR Outcome Evaluation Report Office of Research*.

Duwe, G. (2013). The Development, Validity, and Reliability of the Minnesota Screening Tool Assessing Recidivism Risk (MnSTARR). *Criminal Justice Policy Review*. doi: 10.1177/0887403413478821

Fujikawa, J., Kida, T., & Katoh, T. (2011). Extracting refrained phrases from music signals using a frequent episode pattern mining algorithm. *IEEE International Conference on Granular Computing*.

Janzen, C. A., Deokar, A. V., & El-Gayar, O. F. (2013a). Discovering Predictive Event Sequences in Criminal Careers. *Annual Symposium on Information Assurance (ASIA)*. Albany, NY.

Janzen, C. A., Deokar, A. V., & El-Gayar, O. F. (2013b). Non-parametric discovery of event sequence patterns in criminal behavior. *Proceedings of the 46th Annual Hawaii International Conference on Systems Science (HICSS-46 '13) Symposium on Credibility Assessment and Information Quality in Government and Business*. Maui, HI: IEEE Computer Society.

Keogh, E., & Lin, J. (2004). Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowledge and Information Systems, 8*, 154-177. doi: 10.1007/s10115-004-0172-7

Leleu, M., Rigotti, C., Boulicaut, J. F., & Euvrard, G. (2003). Constraint-based mining of sequential patterns over datasets with consecutive repetitions. *Knowledge Discovery in Databases: PKDD 2003*, 303-314. doi: 10.1.1.134.3862

Martin, B., & Dine, S. V. (2008). Examining the Impact of Ohio's Progressive Sanction Grid, Final Report.

Min, Y., Wong, S. C. P., & Coid, J. (2010). The Efficacy of Violence Prediction: A Meta-Analytic Comparison of Nine Risk Assessment Tools. *Psychological bulletin, 136*, 740-767.

Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poisson model. *Criminology, 31*, 327-362.

Newman, W. J., & Scott, C. L. (2012). Brown v. Plata: prison overcrowding in California. *The journal of the American Academy of Psychiatry and the Law, 40*, 547-552.

Turner, S., Braithwaite, H., Kearney, L., Murphy, A., & Haerle, D. (2012). Evaluation of the California Parole Violation Decision-Making Instrument (PVDMI). *Journal of Crime and Justice, 35*, 269-295. doi: 10.1080/0735648X.2012.683636

Criminal Careers, Criminal Violence, and Substance Abuse in California, 1963-1983, Inter-university Consortium for Political and Social Research (ICPSR) [distributor] (2006).

Zaki, M. J., Lesh, N., & Ogihara, M. (1998). PLANMINE: Sequence Mining for Plan Failures. *4th Intl. Conf. Knowledge Discovery and Data Mining*.