Adaptive Fraud Detection Using Benford's Law

Fletcher Lu[1], J. Efrim Boritz[2], and Dominic Covvey[2]

[1] Canadian Institute of Chartered Accountants,
66 Grace Street, Scarborough, Ontario, M1J 3K9
f2lu@ai.uwaterloo.ca
[2] University of Waterloo, 200 University Avenue West,
Waterloo, Ontario, Canada, N2L 3G1
jeboritz@watarts.uwaterloo.ca,
dcovvey@csg.uwaterloo.ca

Abstract. Adaptive Benford's Law [1] is a digital analysis technique that specifies the probabilistic distribution of digits for many commonly occurring phenomena, even for incomplete data records. We combine this digital analysis technique with a reinforcement learning technique to create a new fraud discovery approach. When applied to records of naturally occurring phenomena, our adaptive fraud detection method uses deviations from the expected Benford's Law distributions as an indicators of anomalous behaviour that are strong indicators of fraud. Through the exploration component of our reinforcement learning method we search for the underlying attributes producing the anomalous behaviour. In a blind test of our approach, using real health and auto insurance data, our Adaptive Fraud Detection method successfully identified actual fraudsters among the test data.

1 Introduction

In this paper we illustrate the implementation of a fraud discovery system which uses a new approach for discovering fraud that combines a reinforcement learning (RL) technique with a digital analysis method. The idea behind this approach is to use the digital analysis to uncover data anomalies and then utilize the reinforcement learning component to reveal the attributes contributing to the anomaly, thereby uncovering underlying fraudulent behaviour.

As Bolton and Hand [2] noted, fraud detection methods may be divided into both supervised and unsupervised methods. For supervised methods, both fraudulent and non-fraudulent records are used to train a system, which then searches and classifies new records according to the trained patterns. The limitation to supervised methods is that one must have both classes of records identified for the system to train on. Thus, this approach is limited to only previously known methods of committing fraud.

Unsupervised methods, in contrast, typically identify records that do not fit expected norms. The advantage of this approach is that one may identify new instances of fraud. The common approach to this method is to use forms of outlier detection. The main limit to this approach is that we are essentially identifying anomalies that may or may not be fraudulent behaviour. Just because a behaviour is anomalous does not necessarily mean that the behaviour is fraudulent. Instead they can be used as indicators of

L. Lamontagne and M. Marchand (Eds.): Canadian AI 2006, LNAI 4013, pp. 347–358, 2006.

possible fraud, whereby the strength of the anomalous behaviour (how much it deviates from expected norms), may be used as a measure of ones confidence in how likely the behaviour may be fraudulent. Investigators may then be employed to analyze these anomalies. However, given the often enormous numbers of records involved in typical fraud application areas such as credit card, cellular phone and healthcare records, even with identified anomalies, investigating the anomalies can be a considerable burden to resources. The novelty of our approach is that we extend the typical outlier detection methods with a reinforcement learning component.

The reinforcement learning component of our algorithm builds on identified outliers by associating with our outliers, underlying attributes that may be linked together to build a case for fraud. Reinforcement learning has typically been used in the past to find the best choice of actions when trying to perform some physical task requiring a sequence of actions to accomplish a desirable goal such as navigating through a maze. The core idea which makes reinforcement learning useful to fraud discovery is its ability to link together states through a pattern of state-action pairs in a policy. In our algorithm we will link together anomalies according to their underlying attributes using the magnitude of the anomalous behaviour as a measure of its desireablity within the reinforcement learning context (in other words anomalies are equivalent to the rewards in an RL environment).

To identify our outliers, we use a digital analysis technique known as adaptive Benford's Law [1]. Benford's Law specifies the distribution of the digits for *naturally* occurring phenomena. For a long time this technique, commonly used in areas of taxation and accounting, was considered mostly a mathematical curiosity as it described the frequency with which individual and sets of digits for naturally growing phenomena such as population measures should appear [3]. Such naturally growing phenomena, however, has been shown to include practical areas such as spending records and stock market values [4]. One of the limits to the use of classic Benford's Law in fraud detection has been its requirement that analyzed records have no artificial cutoffs. In other words, records must be complete. However, in many practical application areas, one only has information for a subset, such as a single year, of financial records. Recent work by Lu and Boritz [1] has removed this limitation with an *adaptive* Benford's Law, making it more practically useful.

In our paper we will explain the algorithm for our technique and test our new fraud discovery technique against outlier detection methods using real healthcare and auto insurance records, demonstrating improvements in classifying fraudsters.

1.1 Complications in Practical Fraud Detection Research

Two major complications for fraud detection researchers are:

1. Secrecy with regards to details on fraud detection techniques.
2. Limits on available real fraud records.

Both of these complications stem from the nature of the application area. It is quite natural that in order to stay ahead of fraudsters, those employing fraud detection methods tend to keep secret their algorithm details in order to avoid fraudsters from knowing these details and developing methods to circumnavigate them. This secrecy makes it

difficult to compare new fraud detection techniques with previous methods. The second complication is due to the fact that companies do not generally like to reveal the amounts of fraud within their field as it tends to have a detrimental impact on shareholder confidence as well as consumer confidence. In addition, laws in Canada as well as many other countries do not require the explicit reporting of fraud losses. Without available real fraud records, many researchers use artificially created synthetic fraud records to test their methods. However, synthetic records for testing are only as useful as they are truly representative of actual fraudulent and non-fraudulent records.

We deal with the first complication by developing a general technique that is not meant to be application area specific but is meant as an improvement over the general Benford's Law outlier detection approach. Many fraud detection techniques are *ad hoc* and use specific details about their application area for detection. For instance, cellular phone detection uses the fact that one can uncover phone usage stealing by identifying that the same phone is calling from two different locations at the same time, which would imply at least one of the calls is from an illegitimate user. Since our fraud discovery technique is meant as an improvement over previous Benford's Law outlier detection methods, we compare our method specifically with that outlier detection approach.

We deal with the second complication by testing our method on real auto insurance records that have been audited and classified for actual fraud. In order to demonstrate robustness, we also apply our method to real healthcare insurance records. However, these records have yet to be audited for fraud and thus we use it only to demonstrate the technique's operation on a differing application area and to illustrate the adaptive Benford's Law component of our fraud detection method on a real fraud application.

2 Background

2.1 Benford's Law

Benford's Law is a probability distribution with strong relevance to accounting fraud. Much of the research on Benford's Law has been in areas of statistics [5, 6] as well as auditing [7, 8].

Benford's Law is a mathematical formula that specifies the probability of leading digit sequences appearing in a set of data. What we mean by *leading digit sequences* is best illustrated through an example. Consider the set of data

$$S = \{231, 432, 1, 23, 634, 23, 1, 634, 2, 23, 34, 1232\}.$$

There are twelve data entries in set S. The digit sequence '23' appears as a leading digit sequence (i.e. in the first and second position) 4 times. Therefore, the probability of the first two digits being '23' is $\frac{4}{9} \approx 0.44$. The probability is computed out of 9 because only 9 entries have at least 2 digit positions. Entries with less than the number of digits being analyzed are not included in the probability computation.

The actual mathematical formula of Benford's law is:

$$P(D = d) = \log_{10}(1 + \frac{1}{d}), \tag{1}$$

where $P(D = d)$ is the probability of observing the digit sequence d in the first 'y' digits and where d is a sequence of 'y' digits. For instance, Benford's Law would state that the probability that the first digit in a data set is '3' would be $\log_{10}(1+\frac{1}{3})$. Similarly, the probability that the first 3 digits of the data set are '238', would be $\log_{10}(1+\frac{1}{238})$. The numbers '238' and '23885' would be instances of the first three digits being '238'. However this probability would not include the occurrence '3238', as '238' is not the *first* three digits in this instance.

2.2 Benford's Law Requirements

In order to apply equation 1 as a test for a data set's digit frequencies, Benford's Law requires that:

1. The entries in a data set should record values of similar phenomena. In other words, the recorded data cannot include entries from two different phenomena such as both census population records and dental measurements.
2. There should be no built-in minimum or maximum values in the data set. In other words, the records for the phenomena must be complete, with no artificial start value or ending cutoff value.
3. The data set should not be made up of assigned numbers, such as phone numbers.
4. The data set should have more small value entries than large value entries.

Further details on these rules may be found in [3]. Under these conditions, Benford noted that the data for such sets, when placed in ascending order, often follows a geometric growth pattern.[1] Under such a situation, equation 1 specifies the probability of observing specific leading digit sequences for such a data set.

The intuitive reasoning behind the geometric growth of Benford's Law is based on the notion that for low values it takes more time for some event to increase by 100% from '1' to '2' than it does to increase by 50% from '2' to '3'. Thus, when recording numerical information at regular intervals, one often observes low digits much more frequently than higher digits, usually decreasing geometrically.

Adaptive Benford's Law modifies classic Benford's Law by removing the second requirement of 'no built-in minimum or maximum values', thus allowing for the technique to be more generally applicable to a wider array of real records which often are incomplete. For more details on the Adaptive Benford method see [1].

2.3 Reinforcement Learning

In reinforcement learning, an environment is modelled as a network of states, $\{s \in S\}$. Each state is associated with a set of possible actions, $a_s \in A_s$ and a reward for entering that state $\{r_s \in R_s\}$. We can transition from one state s(i) to another s(j) by choosing an action a_s and with a certain probability $P(s_j|s_i, a_{s_i})$ we transition to another state. A policy is a mapping of states to action. The objective is to find an optimal policy that maximizies the long-term rewards one may obtain as one navigates through the network. In order to find an optimal policy, we perform a task known as policy evaluation

[1] Note: The actual data does *not* have to be recorded in ascending order. This ordering is merely an illustrative tool to understand the intuitive reasoning for Benford's law.

which determines value estimates for states given a fixed policy, π. The value estimate for a state represents the sum of the discounted future rewards for a state following a policy π. A variety of methods for policy evaluation have developed over time such as the maximum likelihood approach[9], the temporal differencing method [10] and the monte carlo matrix inversion method [11].

One property of the reinforcement learning approach is that it is designed to handle intractably large state spaces. This trait makes it well suited to an environment such as fraud detection where there are usually extremely large numbers of records to process in order to find the relatively small number of occurences of fraud among the total amount of data available. Reinforcement learning also incorporates an exploration component that allows it to search for better actions leading to higher long-term reward values. Such an exploration component is key also for finding *new* instances of previously unknown fraud cases as we wish our fraud discovery method to be able to do. These two traits are the main components motivating our use of a reinforcement learning approach for fraud detection.

3 Algorithm

As we noted in the introduction, our fraud detection method's objective is to improve over outlier detection methods for finding instances of fraud. Outlier detection methods, as we stated previously, actually only indicate anomalous instances. In order to determine whether an anomalous instance is actually a result of a fraudulent occurence typically requires the careful analysis of an auditor. Therefore, we need to consider how an auditor actually determines fraud. Without domain specific knowledge, such as the cell phone example we gave in the introduction, one typically builds a case for fraud by linking together suspicious occurences. Even with domain specific knowledge, such as our cellular phone example, one may still need to link suspicious (anomalous) cases together. For example, even in the cellular phone example where we know that the same cell phone cannot be calling from two different locations at the same time, we do not necessarily know which of the two calls is the fraudulent one and indeed both may be from illegal users. Thus, what is needed is to build a case of fraud by linking together anomalous instances that are related by some set of traits.

Reinforcement learning is well suited to linking together states through its state-action policy mapping. For reinforcement learning to be used as a fraud detection method, we need to be able to relate rewards with anomalies. We do so by setting the reward values as the magnitude that our anomalous instances deviate from expected values. In turn, the states of our RL environment relate to individual records of our application environment and the actions are the attributes of a record. In such a way, two records, with the same attributes are linked together by a common attribute just as an action can relate two states of a classic reinforcement learning environment network.

The best way to illustrate our fraud detection algorithm is through an example. Figure 1 is an example of purchase records for some consumer. Included in the record are the purchased item, the store it was purchased in, the location of the store, the form of payment used to buy the item and the amount of the purchase under 'Digit Sequences'. We apply our fraud detection algorithm by first determining if there are

States	Actions/Attributes				Digit
	Purchase Item	Store	Location	Form of Payment	Sequences
1	shoes	storeA	street15	credit	$52
2	hat	storeB	street12	cash	$38
3	hat	storeC	street17	debit	$22
4	TV	storeB	street11	cheque	$640

Fig. 1. Sample Application: Purchase Records

States	Actions/Attributes				Digit	Rewards/Magnitude
	Purchase Item	Store	Location	Form of Payment	Sequences	of Anomalies
1	shoes	storeA	street15	credit	$52	1.6
2	hat	storeB	street12	cash	$38	3.2
3	hat	storeC	street17	debit	$22	6.2
4	TV	storeB	street11	cheque	$640	1.1

Fig. 2. Sample Application: Analysing Digits with Rewards

any sets of digit sequences that conform to a Benford distribution. In our example there is only one set of numbers, the purchase values that can be compared with a Benford distribution. We compute the frequency with which each digit sequence from 1 to 999 appears in our purchase value records.[2] We then compare these actual digit frequencies with Benford's Law's expected frequencies. If they fall within a certain confidence interval, we will accept that the numerical data follows a Benford's Law distribution.

Assuming the purchase records are a Benford distribution, then we compute a measure of how much any given purchase value deviates from expected Benford value by:

$$Reward(i) = \frac{f_{1i}}{b_{1i}} + \frac{f_{2i}}{b_{2i}} + \frac{f_{3i}}{b_{3i}}, \tag{2}$$

where f_{ji} is the frequency that a digit sequence of length j for state i appears in the dataset and b_{ji} is the expected Benford's Law distribution frequency that the digit sequence of length j for state i should appear.

As an example, consider figure 2 where state 2 has a purchase value of $38. In our algorithm we consider only leading digit sequences.[3] Therefore there are two leading digit sequences, the sequence of just '3' as well as '38'. If 3 appears 22 times in our record, while 38 appears 5 times, then $f_{12} = 22$ and $f_{22} = 5$. Assuming that Benford's Law states that the sequence '3' should appear 10 times and '38' should appear 5 times, then $b_{12} = 10$ and $b_{22} = 5$. Note since there are not three digits in our purchase value $f_{33} = 0$ and does not contribute to the reward function of equation 2. We thus produce a Reward value for state 2 of $Reward(2) = \frac{22}{10} + \frac{5}{5} = 3.2$. We thus can compute reward values associated with each state. Figure 2 illustrates our records with their associated computed rewards.

[2] Benford's Law works with digit sequences of any length. For most practical purposes, the frequencies of sequences of three digits or less are evaluated. For longer digit lengths, the probabilities become so small that they are of little practical value.

[3] See [1] for a justification for this choice.

States	Actions/Attributes				Digit	Rewards/Magnitude
	Purchase Item	Store	Location	Form of Payment	Sequences	of Anomalies
1	shoes	storeA	street15	credit	$52	1.6
2	hat	storeB	street12	cash	$38	3.2
3	hat	storeC	street17	debit	$22	6.2
4	TV	storeB	street11	cheque	$640	1.1

Fig. 3. Sample Application: Choosing a Record/State

Once the reward values have been computed, we can now explore our environment as a reinforcement learning network. We do so by first choosing a start state. In figure 3 we chose state 2. This results in a reward value of 3.2. We then need to choose an action. Our actions are any unused attributes of our record. In this case we have four possible actions. There are numerous methods for choosing an action. See [9] for various techniques.

States	Actions/Attributes				Digit	Rewards/Magnitude
	Purchase Item	Store	Location	Form of Payment	Sequences	of Anomalies
1	shoes	storeA	street15	credit	$52	1.6
2	hat	storeB	street12	cash	$38	3.2
3	hat	storeC	street17	debit	$22	6.2
4	TV	storeB	street11	cheque	$640	1.1

Fig. 4. Sample Application: State to Action to Next State Transition

If we choose action/attribute 'Store'. The specific instance of this action in state 2 is 'storeB'. We therefore search the store column for any other states/records with 'storeB' as an entry. Every possible record with such an entry is a possible next state. In our example state 4 is a possible next state which, as figure 4 illustrates will be our next state. We use a uniform random distribution to choose which of our possible next state candidates will be selected.

Now that we have a method of exploring our environment, we can apply a reinforcement learning algorithm such as temporal differencing or maximum likelihood to find an optimal policy to our system. Any such policy will link together records with the greatest anomalies forming a pattern that builds a case of fraudulent activity just as an auditor may do.

Once you have an optimal policy, the states/records that the auditor wishes to build a case for fraud for, may be used as the start state and then the auditor simply executes the optimal policy choices to find all states that are most strongly linked to that given start state. In this way, the auditor finds all underlying attributes/actions that are in common with the high reward returning states. This requires only a single trace through the system since the optimal policy has already done the exploration that an auditor would have traditionally had to do saving time and man-hours.

A few details which we have not gone into due to space constraints include how to choose when to stop your explorations, handling records that contain multiple columns with digits conforming to a Benford distribution as well as dealing with exploring non-optimal policies. An important reason to obtain less than optimal policies is that less

than optimal policies may still also contain fraudulent patterns. One method to obtain such less than optimal policies is by iteratively removing the optimal policy and then rerunning the algorithm.

In addition, two points for those unfamiliar with Benford's Law and auditing in general may question are that the reward structure of equation 2 does not give higher rewards for digits that appear less frequently than the Benford's Law predicts and that numbers with more digits will possibly have a bias to higher rewards since there are more digits that can contribute to the reward value. Regarding the first point, one could use an alternative formula to equation 2 such as:

$$Reward(i) = q_{1i} + q_{2i} + q_{3i}, \tag{3}$$

where

$$q_{ji} = \begin{cases} \frac{f_{ji}}{b_{ji}} & \text{for } f_{ji} > b_{ji} \\ \frac{b_{ji}}{j_{ji}} & \text{otherwise} \end{cases}. \tag{4}$$

We use equation 2 in our implementation due to our application area of insurance where generally, auditors are most interested in over-charges rather than under-charges. Naturally, this may not be the case in other application areas such as tax filings where one may be concerned with under reporting of figures. In regards to the second point about numbers with more digits having higher rewards, this situation is a Benford's Law phenomenon. Certainly more digits provide more information to the auditor, but this bias is generally not of concern in practical auditing terms as successive digits have geometrically (10 times less for ever digit position to the left you move) less influence on the over all reward relative to the higher leading digits. Further discussions on the influence of successive significant digits may be found in [3]. The authors will gladly provide further details on any of these points on request.

4 Experiments

As stated in section 3 our application analyzes the first, the first two and the first three digit sequences of a dataset, comparing them to expect Benford frequencies. Therefore, for each data set analyzed, we are comparing it to three different Benford's Law digit frequency distributions, one of only length 1 digit, another of length 2 digits and a third of length 3 digits. For conciseness we have reported our digit frequency analyzes through graphs whereby we include the first digit frequencies 1 to 9 on the x-axis values from 1 to 9, the first two digit sequences composed of 10, 11,...,99 on the x-axis from 10 to 99 and the first three digit sequences on the x-axis from 100,...999. When looking at the numerical results, one should keep in mind a single graph is actually three sets of different distributions. The expected Benford curve will contain disjunction points at the points between 9 to 10, and 99 to 100 because they are the points at which a new Benford probability distribution starts and ends.

As stated in section 1.1, we have two databases with which to test our fraud detection method on, a record of healthcare insurance claims provided by Manulife Financial as well as records of auto insurance claims. However, the healthcare insurance claims have yet to be audited and therefore we use it only to illustrate both the robustness of how

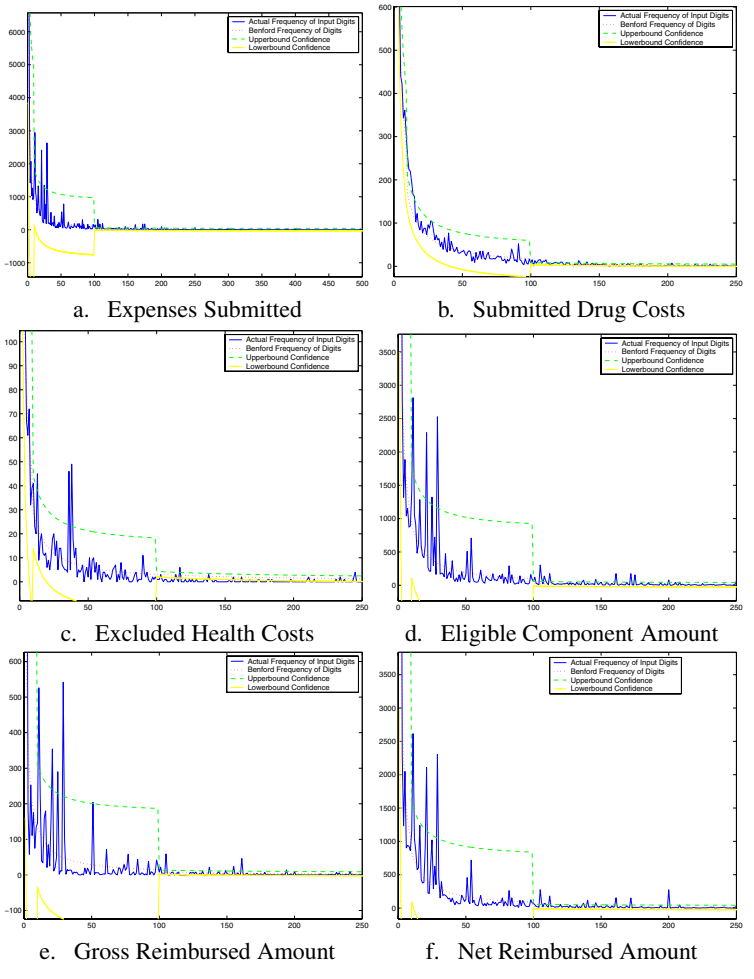

Fig. 5. Healthcare Digit Frequencies compared with their Benford Distributions

our system is generalized enough to operate on differing application areas as well as the operation of identifying anomalous data in conjunction with the attributes related to those anomalies. The second database composed of auto insurance has been audited with identified fraudulent and non-fraudulent records with associated attributes. We therefore use the second database as a test of the utility of our system for identifying real fraud cases.

The healthcare data consisted of 94 columns/attributes with 31,804 rows/records. We therefore had 94 candidate data sets to test to see if they conform to a Benford distribution. 83 of the columns were eliminated due to one of the three Adaptive Benford's Law requirements not being satisfied. Of the remaining 11 data sets, 5 were eliminated using a 90% confidence interval based on training data provided by Manulife. Figure 5 illustrates the frequencies of the digit sequences for the remaining six data sets

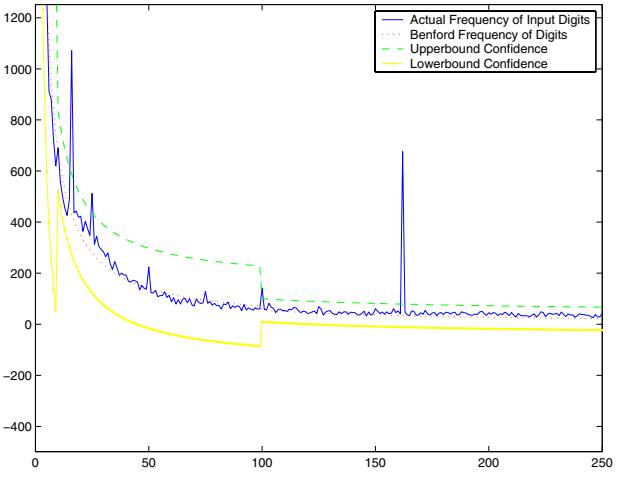

Fig. 6. Auto Insurance Digit Frequencies

that do conform to the Benford distributions. Included on the graphs are the 90% lower and upper bound confidence intervals as well as the the expected Benford frequencies.[4] The six graphs from a to e of figure 5 resulted in 23, 31, 8,21, 14 and 18 respectively of total suspicious attributes, which are currently being investigated for by auditors for fraud.

For our auto insurance database, Ernst and Young provided data which was already audited for fraud. The researchers conducted a test of the system as a blind test where we were unaware of which, if any, of the records and attributes were fraudulent. In this case, only one column of the database conformed to a Benford distribution satisfying the three requirements of Adaptive Benford. Figure 6 illustrates the digit frequencies of the auto insurance data. There were 30 remaining non-Benford attribute columns which may be used as actions to match on in our Reinforcement Learning environment. The database consisted of 17,640 records.

Figure 7 illustrates corresponding rewards resulting from our most optimal policy to our fourth most optimal policy. We obtained the most optimal, second most optimal, etc. polices by successively eliminating the optimal policy currently found and rerunning the optimal policy search. The thirty columns of our graphs represent the thirty attributes that are possible to be chosen from. The height of the bars is the sum of the reward values that would be obtained following the respective policy for each type of attribute the policy chooses. As one can see, the heights for the successively worsening policies result in, not surprisingly, worsening reward values.

Our optimal policy successfully identified the one company that was producing several fraudulent insurance claims. This single fraud generating company corresponds to two of the large spikes appearing in figure 6. However, an important note on our

[4] Confidence intervals were computed based on variance values of training data provided by Manulife and used 2 standard deviations from the expected Benford frequencies for the upper and lower bounds.

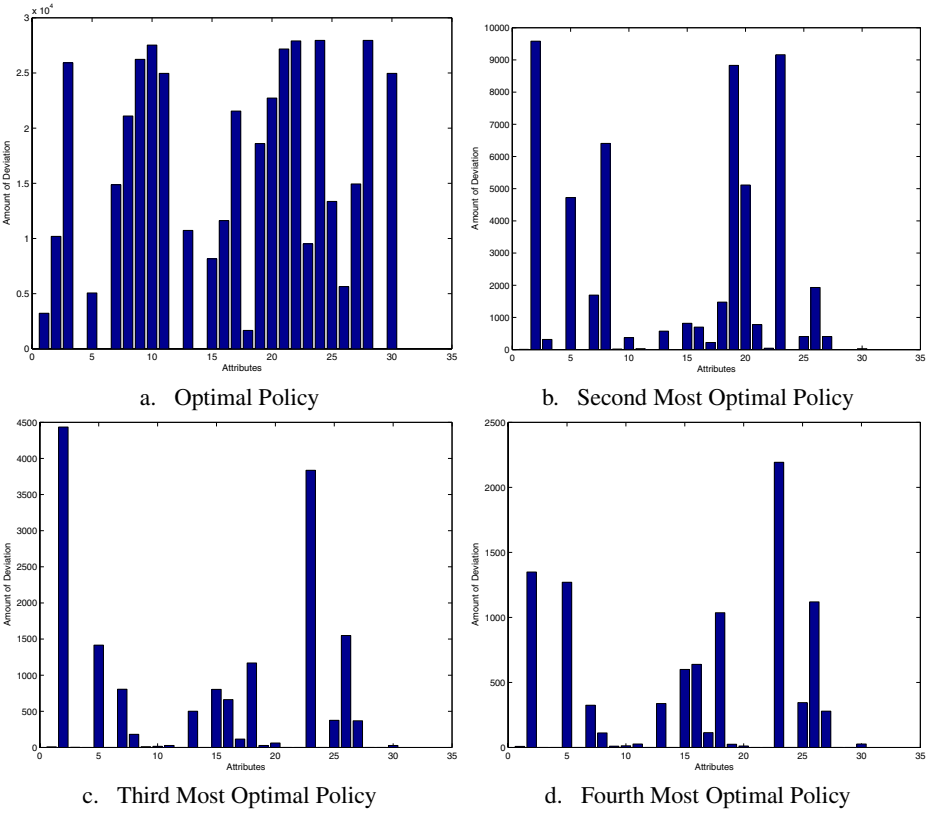

Fig. 7. The attributes of successively worsening policies and the reward values they correspond to across all records

method is that standard Benford outlier methods would have not identified this company because although there was a 'company name' attribute field, this company was not listed consistently in the same way in that field. Instead, they used variations on the name such as abbreviations and reorderings of the words that the name consisted of. Our policy was able to identify the company instead by linking together agents who submitted fraudulent claims on behalf of the company with their locations and in turn to the company. This linking is exactly the kind of fraud case building that the reinforcement learning component is designed to build.

5 Conclusions

In this paper we have presented a new fraud detection method which expands on the current simpler outlier detection approaches. We specifically used a Benford distribution as a benchmark for our unsupervised learning method to discover new fraud cases. We enhanced the method with a reinforcement learning model of our environment in order to link together anomalous outliers to build a case for fraud. In so doing, we are

essentially simulating the behaviour of human auditors. We tested our system with a blind test on auto insurance data successfully identifying instances of fraud perpetrated by several people but linked to one underlying company.

In terms of future research, we plan to incorporate other outlier detection methods with our current Benford method as well as apply our method to greater amounts of audited data from a variety of different application areas.

Acknowledgements. The authors wish to thank Manulife Financial as well as Ernst and Young for providing the insurance data. We would also like to thank the Natural Sciences and Engineering Research Council (NSERC) for providing funding.

References

1. F. Lu and J. E. Boritz. Detecting Fraud in Health Insurance Data: Learning to Model Incomplete Benford's Law Distributions. In *16th European Conference on Machine Learning*, pages 633–640, Porto, Portugal, 2005. Springer.
2. Richard J. Bolton and David J. Hand. Statistical Fraud Detection: A Review. *Statistical Science*, 17(3):235–255, 1999.
3. Mark J. Nigrini. *Digital Analysis Using Benford's Law*. Global Audit Publications, Vancouver, B.C., Canada, 2000.
4. Mark J. Nigrini. Can Benford's Law Be Used In Forensic Accounting? *The Balance Sheet*, June:7–8, 1993.
5. Roger S. Pinkham. On the Distribution of First Significant Digits. *Annals of Mathematical Statistics*, 32:1223–1230, 1961.
6. Theodore P. Hill. A Statistical Derivation of the Significant-Digit Law. *Statistical Science*, 4:354–363, 1996.
7. Charles A.P.N. Carslaw. Anomalies in Income Numbers: Evidence of Goal Oriented Behaviour. *The Accounting Review*, 63:321–327, 1988.
8. Nita Crowder. Fraud Detection Techniques. *Internal Auditor*, April:17–20, 1997.
9. R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts, 1998.
10. R. S. Sutton. Learning to predict by the method of Temporal Differences. In *Machine Learning*, volume 3, pages 9–44, 1988.
11. F. Lu and D. Schuurmans. Monte Carlo Matrix Inversion Policy Evaluation. In *UAI: Proceedings of the 19th Conference*, pages 386–393, San Francisco, 2003. Morgan Kaufmann.
12. Mark J. Nigrini and Linda J. Mittermaier. The Use of Benford's Law as an Aid in Analytical Procedures. In *Auditing: A Journal of Practice and Theory*, volume 16(2), pages 52–67, 1997.