Optimizing Debt Collections Using Constrained Reinforcement Learning

Naoki Abe Prem Melville Cezar Pendus David L. Jensen NABE,PMELVIL,CPENDUS,DAVJEN@US.IBM.COM

Business Analytics and Mathematical Sciences Department, IBM Research, Yorktown Heights, NY 10598

Chandan K. Reddy REDDY@CS.WAYNE.EDU

Department of Computer Science, Wayne State University, Detroit, MI 48202

Vince P. Thomas James J. Bennett Gary F. Anderson Brent R. Cooley VPTHOMA, JJBENNET, GFA, BCOOLEY@US.IBM.COM

Global Business Services, IBM Corporation, Armonk, NY 10504

Melissa Weatherwax Timothy Gardinier Gerard Miller MILISSA_WEATHERWAX, TIMOTHY_GARDINIER, GERARD_MILLER@TAX.STATE.NY.US

State of New York, Dept. of Tax. and Fin. W. A. Harriman Campus, Albany, NY 12227

Abstract

The problem of optimally managing the collections process by taxation authorities is one of prime importance, not only for the revenue it brings but also as a means to administer a fair taxing system. The analogous problem of debt collections management in the private sector, such as banks and credit card companies, is also increasingly gaining attention. With the recent successes in the applications of data analytics and optimization to various business areas, the question arises to what extent such collections processes can be improved by use of leading edge data modeling and optimization techniques. In this paper, we propose and develop a novel approach to this problem based on the framework of constrained Markov Decision Process (MDP), and report on our experience in an actual deployment of a tax collections optimization system at New York State Department of Taxation and Finance (NYS DTF).

The tax/debt collections process is complex in nature and its optimal management will need to take into account a variety of considerations. At a high level, the collections process management problem is that of determining the answers to the following questions: (1) which of the debtors should be approached; (2) which of the possible collection actions are to be taken onto them; (3) who should take those actions; and (4) when they should be taken. The answer to each of these questions will depend on a number of factors. The answer to the first question will depend on the information the collection agency may have on the debtors, such as the demographics and the amount of debt owed. The answer to the second will also depend on the nature and status of the debtor, such as how collectible they appear to be, since actions of varying severity may be appropriate for debtors in different categories. The third will additionally depend on what resources are available within the collection organization. With regard to the fourth, there are several complications that impact the optimal timing to take actions on debtors. For example, in the tax collection case, there are certain legal requirements that govern the sequential course of collection actions. The prime example is the requirement that a warrant need be issued before collection actions such as levy and seizure are performed. Also, there are business considerations that affect the appropriate timing of a collections action in general.

Due to the high complexity involved in the collection process, and particularly because of the legal and business constraints, it is

Appearing in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011.

common practice to follow rigid manually constructed rules to guide the collection activities. Even in state-of-the-art rule based systems for collections, the role of data analytics is typically limited to that of augmenting the rules engine with scoring functions that they can refer to. In the present work, we develop a collection process automation system that is *primarily* based on data modeling and optimization, and accepts input from a rules engine as constraints on the optimization process.

Specifically, we propose a novel approach based on the framework of constrained Markov Decision Processes (MDP), and develop concrete algorithms and methods for performing data modeling and resource optimization in a unified fashion, while addressing all of the considerations that were touched upon in the foregoing discussion. The MDP formulation is particularly applicable here because of its treatment of the sequential dependencies between actions via the introduction of states. For example, the legal requirement that a warrant precedes levy and other collection actions can be addressed naturally by introduction of states that correspond to warranted cases and the unwarranted cases. The optimal action in a non-warranted state may very well be to warrant, whereas the optimal action in a warranted state may be to levy. Furthermore, the value of issuing a warrant in an unwarranted state cannot be fairly assessed without considering the long term rewards obtained in future (warranted) states, by subsequent collection actions such as levy. The consideration of long term rewards as the objective employed in MDP is therefore essential for satisfactory formulation of the collection process.

The constrained MDP formulation, beyond the standard MDP, allows us to take into account the various constraints that govern the actions under consideration. For example, the value of a warrant now will likely depend on the available resources for performing subsequent corresponding actions such as levy. Estimating the value of an action in a constrained MDP involves looking ahead into the resources available at the future states, and the resulting value function and corresponding optimized policy will be better guided with respect to resource consumptions. In addition to the resource constraints, the business and legal constraints can also be handled by including them as an additional type of constraints that define the set of legal policies in the constrained MDP.

As described above, the proposed approach based on the constrained MDP framework provides a comprehensive solution to the problem of collections optimization, which tightly couples data modeling and constrained optimization in a unified manner. Furthermore, it can accept input from rules engine as constraints, and the output policy can be fed into a rules engine. The approach thus allows us to invert the typical roles played by the rules engine and analytics/optimization engine, and provide a nearly automated system in which rules are fed as input to and output from the analytics/optimization engine.

The system became operational in December 2009, and the results to date are encouraging. New York State has observed an $83 million increase in revenue from 2009 to 2010 (8%), using the same set of resources. Given a typical annual increase of 2-4%, the expected benefit of the developed system is approximately $120 to $150 million over the next three years, far exceeding the initial target of $90 million. It is reasonable to expect this to improve in the future as the system further adapts. We have also seen improved productivity in the primary enforcement actions taken by the tax department. Specifically, the dollars collected per warrant has increased by 22%, while the number of warrants went down by 9%. The dollars per levy went up by 11%, and the number of levies served decreased by 3%. As a result, 35 thousand less taxpayers had these serious enforcement actions taken against them - a win for all.