Pergamon

Mathl. Comput. Modelling Vol. 19, No. 1, pp. 1–6, 1994

0895-7177/94 $6.00 + 0.00

0895-7177(94)E0001-4

The Theoretical Model for the Abandoned Mines Budget Allocation

W. HOLSZTYŃSKI* AND W. W. KOCZKODAJ†
Canadian Institute for Mining and Mineral Institute
Computer Science, Laurentian University, Sudbury, Ontario, Canada P3E 2C6

(Received and accepted December 1993)

Abstract—The government assigns a budget for rehabilitating abandoned mines. While it is usually impossible to accomplish a total rehabilitation within a limited budget, we will show how to achieve the best possible allocation of funds.

We present our solution in two stages. First, we describe a simple solution which is already a good approximation to the best solution. Then we outline a solution which is the best within the precision of any expert's ability to judge the cost and the benefit of the undertaken remedial actions.

1. PROBLEM DEFINITION

Any decision related to abandoned mine hazards must contribute to the safe and secure living conditions of current and future generations. The best solution is to remove all symptoms of the abandoned mine which might be hazardous for people and the environment. This, however, is unrealistic in a short term perspective. Limited public and private funds should be allocated in a manner that will improve the public health conditions and safety.

Each rehabilitation remedial action should have an effect which we call benefit. Our guiding principle is to achieve the maximal possible benefit per dollar spent from a budget.

The complication inherent in the considered situation is due to the fact that remedial actions are interrelated. A rehabilitation of a mine may consist of several remedial actions such that the sum of the costs of fulfilling them separately might be higher than the cost of doing all of them at one time. Furthermore, the benefit from accomplishing each of the interrelated remedial actions might be higher if other remedial actions are accomplished too.

Nevertheless, for the sake of simplicity, first we consider the problem under the assumption that the costs and the benefits from accomplishing different remedial actions are independent. It is worthwhile to implement this approach all the way down to a software system. The software simulation will serve to get a clear idea of the situation at hand by providing a rough solution to the problem of mine rehabilitation. It will also serve to test and analyze the final, more refined, solutions. We will call this simpler model "the independent remedial actions case."

The main conceptual notions which we will use are: remedial action, cost (of the remedial action), benefit (from accomplishing a remedial action), the budget, the total cost of a solution (which must not exceed the budget), and the total benefit of a solution (which should be maximal). In the advanced analysis, the cost and the benefit will be associated with some sets of interrelated remedial actions. This may lead to insurmountable complications. However, we reformulate the general case in such a way that the issue of interrelation is replaced altogether by a simpler

*An independent contractor presently working at Wiltel in The Woodlands (Texas) subdivision.
†Partially supported by the Natural Science and Engineering Council of Canada under Grant OGP 003838.

notion of disjoint clusters of remedial actions. Hence, we will be able to solve the general problem efficiently (see Section 3) in a way which takes into account advantages related to the proximity of the actions. It is reasonable to assume that a cluster may be associated with a mine site on which there are number of hazards.

The problem at hand is of a finite discrete nature (since the number of mines and remedial actions is finite), but to show clearly the direction of our ideas to the reader, first we will briefly describe the continuous version of the independent remedial actions case in Section 2. Then in Section 3, we will write about the algorithmic approaches to the case of independent remedial actions. Section 3 is devoted to the general case.

REMARK. Some readers may prefer to skip Section 2 (especially if they are not fond of the measure theory) and start their reading from Section 3.

2. CONTINUOUS MODEL

When the number of remedial actions is high while the cost of each remedial action is a comparatively small fraction of the budget, we encounter a situation which is approximated well by a continuous model. This is one more reason for considering the continuous model.

The model consists of an atom free measure space of remedial actions (T, cost) and of a measurable nonnegative real-valued function $p : T \to R$, which is a derivative of the benefit with respect to the cost in the following sense.

Let A be a measurable subset of T (i.e., $\text{cost}(A)$ is defined). Then

$$\text{Bene}(A) = \int_A p \, d(\text{cost}).$$

The above equation serves as a definition of the benefit function Bene in our continuous model.

The goal is to find a measurable subset S of T such that under the constrain $\text{cost}(S) \leq \text{budget}$, $\text{Bene}(S)$ is maximal—then S is the optimal set of remedies; in general, such an optimal set does not need to be unique.

We make realistic assumptions that $\text{cost}(T)$ and $\text{Bene}(T)$ are finite.

Furthermore, if $\text{cost}(T)$ does not exceed the budget, then the problem is trivial: $S = T$ is obviously the (most satisfactory!) solution. Unfortunately, the available funds are seldom sufficient to cover all desirable remedial expenses. Thus, in what follows, we assume (realistically again) that: $\text{cost}(T) > \text{budget}$.

REMARK. We could also start from the benefit function as one of the primary notions from which we would derive the notion of the derivative p of the benefit with respect to cost as in the Radon-Nikodym Theorem (see [1]).

Let s be the greatest lower bound of all real numbers y for which $\text{Bene}(p^{-1}([y, \infty)) < \text{budget}$.

Let $\tilde{S} = p^{-1}([s, \infty))$. Then $\text{cost}(\tilde{S}) \geq \text{budget}$. If $\text{cost}(\tilde{S}) = \text{budget}$, then we let $S = \tilde{S}$. Set S is a solution and even (in this case) the unique solution to our problem.

Otherwise, when $\text{cost}(\tilde{S}) > \text{budget}$, then there exists (not unique) subset A of $p^{-1}(y)$ such that $\text{cost}(A) = \text{cost}(\tilde{S}) - \text{budget}$; such a set A exists because we assumed that cost is a measure free of atoms (see [1]). Now we define $S = \tilde{S} \backslash A$. Then $\text{cost}(S) = \text{budget}$ and set S is a desired optimal solution.

REMARK. While solution in general is not unique, all optimal solutions S are contained in $\tilde{S}$. In practice, from the global point of view, the optimal solutions will tend to differ only marginally.

3. DISCRETE MODEL—THE INDEPENDENT REMEDIAL ACTIONS CASE

Let: T be a finite set—elements of T are called remedial actions, $\text{cost} : T \to R$ be a real-valued function which assumes positive values, $\text{Bene} : T \to R$ be another function, which assumes positive values, and X be a subset of T. Then $\text{cost}(X)$ and $\text{Bene}(X)$ are defined as the sum of the costs $\text{cost}(x)$ and benefits $\text{Bene}(x)$ for all elements $x \in X$.

REMARK. These are exactly the above assumptions of additivity of the cost and the benefit functions that are characteristic for the case of independent remedial actions. We will see, however, that the fully general model of interdependent remedial actions can be formulated in a way which doesn't differ too much from the present case.

If $\text{cost}(T) \leq$ budget then, trivially, $S = T$ is the unique optimal solution. Thus, from now on, we will assume (quite realistically) that $\text{cost}(T) >$ budget.

We define a "benefit per dollar" function $p : T \to R$ as follows:

$$p(t) = \frac{\text{Bene}(t)}{\text{cost}(t)}.$$

Let $n = |T|$ be the number of all remedial actions. Let us order the remedial actions $t_0, t_1, \ldots, t_{n-1}$ according to their benefit per dollar values, in the decreasing order: i.e., $p(t_0) \geq p(t_1) \geq \cdots \geq p(t_{n-1})$.

A strong initial candidate for a solution is the set A defined by the following simple "greedy" algorithm (described below in pseudo C language):

ALGORITHM.

```
A = emtpy;
k = 0;                              // initialization
while((cost(A) < budget) and (k < n))
      { if(cost(A) + cost(t_k) ≤ budget)
            A = A ∪ {t_k};          // include t_k into A
            k++;
      }
```

We would normally obtain from the above procedure a set A of the form: $A = \{t_0, t_1, \ldots, t_{k-1}\}$, where k is the smallest integer for which $\text{cost}(t_0) + \cdots + \text{cost}(t_k) = \text{cost}(A) + \text{cost}(t_k) >$ budget.

In the lucky instance (not very likely) when $\text{cost}(A) =$ budget, such a set $A = \{t_0, \ldots, t_{k-1}\}$ is the unique optimal solution. (But from now on we don't assume that A is necessarily of the form $\{t_0, t_1, \ldots, t_{k-1}\}$.

The solution produced by the greedy algorithm tends to be nearly optimal under wide circumstances. In general, when the number of remedial actions is high while the cost of the remedial actions, relative to the budget, are low, then the set A produced by the greedy algorithm still is a nearly optimal (or even an optimal) solution. Even more generally, set A is a nearly optimal solution when the remaining remedial actions, which form set $T \backslash A$, have costs which are relatively low when compared to budget.

In the future, we will develop other advanced algorithms which will produce nearly optimal solutions S, perhaps superior to the above solution A. But solution A is already rather strong and so simple that first we will always estimate how much there is to gain by applying more sophisticated procedures. On the other hand, the example in the Appendix at the end of this note shows that there are limitations to the greedy algorithm.

We said from the beginning that our main guiding principle is to maximize benefit per budget dollar spent. This would be the whole principle in the continuous case, when the entire budget would be utilized. The maximal utilization of the budget (essential when $\text{cost}(T) >$ budget) is our other, auxiliary principle. Indeed, in the discrete case, the sum of costs of any selection of remedial actions might not add up to the full budget. But we will consider as admissible solutions only those which we call *saturated*. These are the subsets X of T such that $\text{cost}(X) \leq$ budget, while $\text{cost}(Y) >$ budget for any superset Y of X (where Y is also a subset of T).

The solution A, produced by the greedy algorithm described above, is always saturated, hence admissible.

4. DISCRETE MODEL, INCLUDING INTERRELATED REMEDIAL ACTIONS

The challenge of rehabilitating a mine may consist of several remedial actions of which, due to the shortage of funds (budget), only some would be done. If we consider two remedial actions at different mines, especially if the distance between mines is significant, then the cost of performing two such remedial actions is the sum of the costs of the remedial actions considered separately. The same is true for the benefit obtained from performing two such remedial actions.

When two or more remedial actions concern only one mine, or even a cluster of mines situated in the same local area, then the above additive rules concerning their total and separated costs and benefits are no longer true. Thus, we are forced to consider selections of remedial actions instead of individual remedial actions when we discuss their costs and benefits.

Thus, remedial actions are clustered into groups (e.g., mine sites), where there is no interdependence between remedial actions from different groups. On the other hand, a group of (perhaps) interdependent remedial actions admits 2^n possible selections (including the empty and the full selection)—a number which grows fast when n increases. Even for a group of $n = 8$ remedial actions, we get an uncomfortably high number of 256 possible selections, and when $n = 12$, then the number of selections would already be 4096. It's clear that there do not exist the means sufficient to evaluate the costs and benefits of all selections, especially that the number of mines in question is in thousands.

It follows that only some selections of (interrelated) remedial actions will be evaluated. From now on, we will call them *clusters*. These *clusters* (of remedial actions) will be identified by the institution responsible for mines rehabilitation. Thus, in the present general model, our fundamental notions are the following: *remedial actions*, *clusters*, *groups* (of remedial actions), *costs*, *benefits*, and *budget*.

Thus, from now on, we use the word group in a formal way. We assume that each remedial action belongs to one and only one group. Clusters are sets of remedial actions, and each cluster is contained in one (and only one, of course) group, which varies for different clusters. Let C be the set of all clusters. Then cost and benefit are real-valued, positive functions defined in C:

$$\text{cost} : C \to R \quad \text{and} \quad \text{Bene} : C \to R.$$

The set of all groups will be denoted by G. Thus $T = \cup G$, i.e., the set of all remedial actions, T, is a union of all groups.

Some (or even all) clusters may consist of one remedial action only.

We assume that for each group $X \in G$ there exists a set F of clusters which partitions X, i.e., such that every remedial action from X belongs to one and only one of the clusters from F.

Attention: in addition to clusters from a partition F of X, group X will most likely contain a lot of clusters outside F as well; furthermore, group X may admit not just one partition like F, but many of them (our minimal requirement is that there is at least one).

REMARK. The last assumption requires an effort on the part of the institution which carries the rehabilitation of mines—there must be an evaluation effort conducted which would deal with the sufficient number of clusters. But our condition is natural and virtually necessary anyway if a reasonable optimization is expected. Otherwise, some groups would not have a chance to be completely included into the rehabilitation program—a rather undesirable situation.

REMARK. On the other hand it's preferable not to evaluate a selection of remedial actions if complementary evaluations, adding up to a partition of the respective group, are not going to be conducted. Such "noncomplementable" evaluations cannot be utilized very usefully, and they certainly diminish the efficiency of the algorithms which search for optimal solutions to the mine rehabilitation program. Thus, the mine rehabilitating institution may save on the effort of performing evaluations of noncomplementable clusters.

The most significant relation between clusters is disjointness. Indeed, by a solution we will mean a saturated collection of clusters, every two of which are disjoint—"saturated" means that no cluster can be added to the collection without the sum of the costs of the clusters exceeding the budget. Once again, our goal is to obtain a solution for which the benefit per budget dollar spent is the highest (or nearly so).

Once again, our first algorithm will be the "greedy algorithm" similar to the one described in Section 2. However, this time we expect that more advanced algorithms, including algorithms based on the concept of simulated annealing, will be more advantageous, more so than in the simpler case of independent remedial actions, covered in Section 2.

5. CONCLUSIONS

We would like to stress the simplicity of our final approach, when compared with the much more complex problem of dealing with the involved notion of interrelated remedial actions. We reduced the situation to studying disjoint clusters which are just as independent as remedial actions were in Section 2, i.e., the cost and the benefit from accomplishing a collection of pairwise disjoint clusters is the sum of the costs and benefits, respectively, of the clusters belonging to the collection. This reduction makes the general rehabilitation problem algorithmically feasible.

APPENDIX

In practice, the cost of a remedial action or of a cluster is only a small fraction of the budget. Thus, in practice, the simple greedy algorithm will perform quite well. This is true for the mine rehabilitation program as well as for any program which consists of relatively small actions. Only when there are large actions will we truly need algorithms which outperform the greedy algorithm. Below are two examples which illustrate the issue. We assume in these examples that all actions are independent from one another.

EXAMPLE 1. We have only three actions t_0, t_1, t_2. The costs are: $\text{cost}(t_0) = \text{cost}(t_1) = 10$, $\text{cost}(t_2) = 100$, while budget $= 119$. The benefits from performing the given actions are $\text{Bene}(t_0) = \text{Bene}(t_1) = 100$ and $\text{Bene}(t_2) = 999$. Thus, the benefits per dollar are: $p(t_0) = 10$, $p(t_1) = 10$, $p(t_2) = 9.99$.

Let's remember and stress again that the above example is superficial and not characteristic of real situations at all. We provide it to give a fuller understanding of our topic, and also to show the limitations of the greedy approach, especially when someone would like to extend our methods beyond the application presented here.

It follows that the greedy algorithm will find solution $S = \{t_0, t_1\}$ which is saturated and which indeed is even optimal since it gives maximal benefit per dollar. But, perhaps, the situation is not satisfactory. Despite the optimality, the budget is far from being fully utilized. Thus, one may redefine the problem; instead of optimizing the benefit per dollar, one may simply optimize the total benefit. In the latter case, $S' = \{t_0, t_2\}$ would be one of the two optimal solutions.

EXAMPLE 2. The budget is again 119. There are 102 actions: $t_0, t_1, \dots, t_{101}$. Their costs are: $\text{cost}(t_0) = \text{cost}(t_1) = 10$, $\text{cost}(t_2) = 100$, and $\text{cost}(t_3) = \cdots = \text{cost}(t_{101}) = 1$, while $\text{Bene}(t_0) = \text{Bene}(t_1) = 100$, $\text{Bene}(t_2) = 799$, and $\text{Bene}(t_3) = \cdots = \text{Bene}(t_{101}) = 1$.

This time the greedy algorithm provides us with solution $S = \{t_0, t_1, t_3, \dots, t_{101}\}$, which is not optimal. It fits the budget perfectly ($\text{cost}(S) = 119$), but its benefit is only $\text{Bene}(S) = 299$; hence, benefit per dollar is under 3; to be precise, it is equal to 299/119.

This time, the greedy algorithm provides us with the solution $S = \{t_0, t_1, t_3, \dots, t_{101}\}$, i.e., S consists of all actions, but t_2. The solution S fits the budget perfectly ($\text{cost}(S) = 119$), but S is not optimal. Indeed, the benefit per dollar is under 3; to be precise, it is equal to 299/119.

One of the optimal solutions is the following: $S' = \{t_0, t_2, t_3, \ldots, t_{21}\}$. The cost of S' equals the budget again: $\text{cost}(S') = 119$ while the total benefit derived from S' is much higher ($\text{Bene}(S') = 828$) and so is the benefit per dollar; it is nearly 7; to be exact, it is $828/119 = 6.966\ldots$

We are reminded that this happens when there are large actions which provide considerable benefit per dollar. In the case of the mine rehabilitation program, each of the actions costs only a small fraction of the budget; hence, the greedy algorithm will perform quite well.

REFERENCES

1. P.R. Halmos, *Measure Theory*, D. Van Nostrand, New York, (1962).
2. Z. Duszak, W.W. Koczkodaj and W.O. Mackasey, Towards better abandoned mine hazard prioritizing—An expert system approach, In *The Challenge of Integrating Diverse Perspectives in Reclamation, Proceedings of 10th National Meeting of ASSMR*, Spokane, Washington, May 16–19, 1993, pp. 577–589, (1993).