

Optimizing infrastructure network maintenance when benefits are interdependent

Ioannis S. Papadakis¹ and Paul R. Kleindorfer²

¹ LeBow College of Business, Drexel University, Acad. Building 55–228A,
Philadelphia, PA 19104, USA (e-mail: isp@drexel.edu)

² The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6366, USA
(e-mail: kleindorfer@wharton.upenn.edu)

Abstract. Network Topology Dependencies (NTD) are a class of externalities in the maintenance cost structure of infrastructure networks with applications to many network industries, including natural gas and water distribution pipelines. It is shown that the above externalities may be included to infrastructure maintenance decisions, if optimal maintenance is formulated as a Rhys-Balinski selection problem. A unique contribution is that this risk management problem is analyzed from the point of view of integrating quantitative analysis to organizational and inter-organizational decision processes. Hence, the importance of various procedural requirements is established in addition to computational efficiency and numerical accuracy. In particular, the benefits of sensitivity analysis facilitation and of avoiding manipulability are stressed. The proposed solution process achieves all four requirements. Special attention is paid to the role of submodularity and antitone differences in sensitivity analysis.

Keywords: Maintenance planning – Network optimization – Decision processes

1 Introduction

Maintaining networks is a common problem to many industries. Examples include natural gas and oil pipelines, underground cable networks, water mains networks, sewer systems, railroads, and road networks. In all these industries the maintenance cost structure can be seen as having two distinct components.

(1) *Direct costs.* These depend on the technological attributes of the capital equipment to be maintained / renewed. The typical example of direct cost is pipe material cost. Direct costs cannot be shared and are necessary if a part of the network is to be maintained.

Correspondence to: I.S. Papadakis

(2) *Network topology dependencies (NTD)*. Two types of NTD, capturing a wide variety of construction – maintenance costs, are considered here: A) *Contiguity Discounts* are realized when costs – that would otherwise be replicated – are paid once when contiguous sections are maintained at the same time. Examples include parts of the excavation cost for pipelines and parts of horizontal drilling costs. B) *Setup Discounts* are realized when costs may be paid once for a neighborhood of the infrastructure network, independently of how much work is carried out on it. Examples include parts of material handling cost, nuisance costs (from traffic disruption and noise), and construction site setup costs.

The operation of the networks we consider, for instance pipelines, railroads and highways, tends to have benefits that are broadly shared in society and risks burdening mainly residents of locations adjacent to network segments. Experts and lay-people are likely to agree on the concept of network dependencies, even though it is more likely that only the network operating company would be in a position to generate a precise quantitative estimate of dependency magnitudes. This information asymmetry may generate difficult to resolve conflicts between the network operating company deciding on network maintenance and residents in neighborhoods close to network segments or other stakeholders concerned about risk equity.

Consider a network operator interested in finding its annual optimal maintenance plan, where segment risk distribution is according to Figure 1. In Figure 1 segment risk is represented by a risk index measuring the ratio of expected liability for the operator to cost of segment maintenance. This risk index can be seen as the benefit cost ratio for the maintenance of the individual segment. Assuming a linear cost structure where the cost of segment maintenance is proportional to length and segments have equal length, the direct cost of maintenance is equal for all segments in the network. There is a very simple and intuitive process for the selection of the optimal maintenance plan in this case (left side of Fig. 1): start with the highest risk index segment and proceed until all segments with risk index higher than one are scheduled for maintenance. This “greedy” algorithm achieves efficiency and equity simultaneously.

Clearly this is the most cost-efficient maintenance plan from the point of view of the network operator. But whether this plan is equitable or not can be established only after equity is carefully defined. If a perfect system of compensation for injuries to network neighbors exists, then the distribution of residual risk after maintenance loses its importance. It would be unrealistic to assume such a system exists, however, especially in the case of physical injuries and fatalities. Thus, residual risk distribution is expected to be an important part of optimal network maintenance decisions. The “highest risk index first” process has the obvious attribute that risks eliminated are higher than risks accepted during network operation. If continuing network operation generates an essential public good, some residual risk is unavoidable. In this common case, minimizing maximum residual risk will be viewed as an equitable decision process from a network neighbor knowing the risk distribution, but blinded by a “Rawlsian veil of ignorance” unaware of who is exposed to which risk. It is very easy to imagine a “veil of ignorance” exists

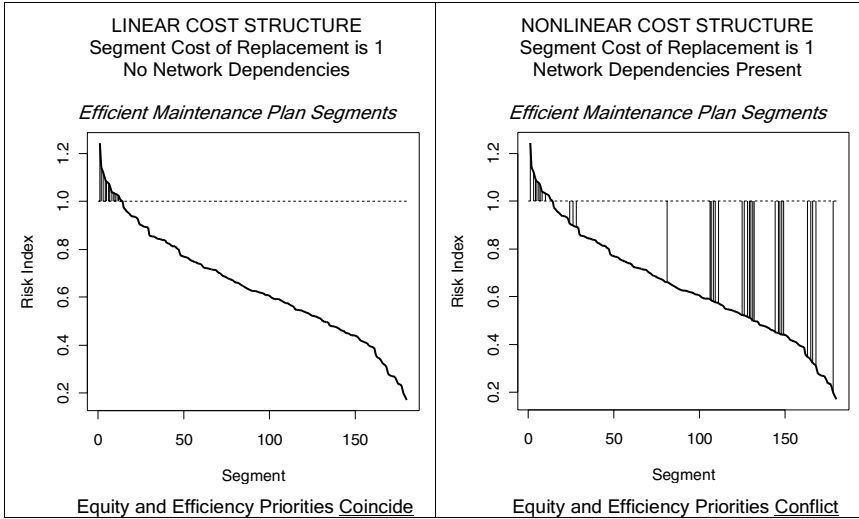


Fig. 1. Risk distribution prior to maintenance with efficient maintenance plan segments marked by a vertical line. Nonlinearities in the cost structure affect the congruence between efficiency and equity priorities

when high-risk locations vary from one decision period to another as well as when network neighbors do not trust the operator.

When network dependencies are significant, the most efficient maintenance plan may look like the one on the right side of Figure 1. Clearly, a nonlinear cost structure may result in a conflict between equity and efficiency priorities. The most efficient maintenance plan accepts high risks and eliminates low risks, thus violating the minimax residual risk attribute of the previous case. Now, the operator may seek to establish that if injuries occur, compensation will be perfect or that inequity in residual risk is justifiable by the significant shared efficiencies from exploiting the nonlinear cost structure (reflected in lower costs of network access). Nonlinear cost structures are associated with more complex decision processes that are not as transparent as the “highest risk first” process. We propose a decision process that is fairly complex, but is verifiable, non-manipulable, and permits maintenance plan overrides to impose equity constraints or management priorities. Hence, we maintain that the proposed decision process may facilitate, not hinder, acceptance of optimal maintenance plans by concerned stakeholders, external to the organization of the network operating company.

Risk assessment, whether it is based on statistical estimation or on risk models, produces estimates within confidence intervals or ranges of uncertainty. It is thus necessary to put risk management decisions, hinging on imprecise risk estimates, under the test of sensitivity analysis. The proposed decision process anticipates and facilitates sensitivity analysis. In particular, we show that the decision model we propose is robust in the sense that it has intuitively appealing properties in response to changes in estimated decision parameters. This is very desirable in avoiding decision process inefficiencies resulting from protracted deliberations and high-

cost data verification, within or across the organizational boundaries of the network operator.

Up to now NTD have received little attention during budgeting. It is clear that in the execution phase of a yearly maintenance plan one can and should avoid replicating certain expenses. But, if NTD are not recognized during budgeting, their value may become deadweight loss. For example, consider a plan that doesn't identify NTD, but assigns all the work in a service subregion to one organizational division. Employees in this division may capture the NTD value by operating at a lower workload. If, however, the division of workload doesn't follow the proper boundaries, then nobody is in a position to capture the value of NTD. A telling example of NTD value loss is having a crew dig a hole on the ground, finish some work, and then cover the hole only to have another crew visit the same area shortly after and redig the same hole. Investor-owned companies and many municipal companies have been operating in the past under a rate-of-return regulation system. It is well known that the efficiency of capital investments doesn't receive the highest priority under cost-plus regulation. However, Crew and Kleindorfer (2002) report that many network industries are switching from rate-of-return regulation to incentive regulation (e.g., price caps) and therefore one can expect more attention to capital investment efficiencies and to NTD in the years to come.

Moreover, institutional innovations in the governance of infrastructure networks (e.g., privatization of network assets) have intensified public interest in the physical risk resulting from the operations of networks and the equity of its distribution. Heuristic solutions for maintenance prioritization – often utilized in practice – are challenged not only on the basis of inefficiency, but also on the basis of undue inequity. Crude heuristics result in inexact solutions that exhibit unpredictable inconsistencies in space and time. In addition to their evident inefficiency, these inconsistencies raise questions about maintenance prioritization being manipulable. The decision process for maintenance prioritization we propose produces exact solutions, hence its results are replicable. We expect that verifiability supports the legitimation of efficient network maintenance even when the resulting residual risk distribution is unbalanced.

The importance of a legitimable and efficient decision model for network maintenance, becomes clearer if one considers the disadvantages arising from its absence. Inefficiency, as a result of failing to recognize the importance of NTD or due to dependence on crude heuristics, forces the network operator to function suboptimally, with lower profitability and higher service price. When a legitimable maintenance decision process is missing, local communities are forced to accept an arbitrary – in their minds – allocation, even though imperfect compensation gives them standing in the risk allocation decision. This sense of injustice results, in turn, to lack of cooperation from local communities and thus to many long term risks for the network company, including unnecessary friction in network expansion projects.

The multiple procedural requirements in the solution of the maintenance problem, legitimation, allocational efficiency, non-manipulability, and decision process efficiency, are met by following an integrative approach. It consists of three equivalent formulations, which are presented in Section 2. First, a quadratic binary

program is presented that applies if infrastructure network topology, section maintenance costs, and expected maintenance benefits are known. This is followed by a Rhys (1970) – Balinski (1970) model, developed for the analysis of investments with mutually exclusive costs and dependencies in benefits. Examples of the latter problem are the “provisioning problem” described by Lawler (2001) and Mamer and Smith’s (1982) “repair kits.” Finally, an equivalent maximum flow – minimum cut model is developed in order for an exact solution to be efficiently obtained. In addition, Section 2 includes a subsection that shows how requiring the maintenance of certain segments can be handled efficiently within the proposed solution framework.

In Section 3 submodularity and antitone difference properties are verified to hold for the problem in hand. Using Topkis’ (1978) work, it is shown that the optimal set of arcs to be maintained using the most conservative estimates will remain in the optimal set if benefits are revised upwards. The latter property facilitates sensitivity analysis in response to often occurring changes in benefit parameters. Furthermore, this property can be utilized to obtain the Benefit-Cost efficiency frontier for this problem quite easily, an essential part of a budgeting exercise. Section 4 includes a complete analysis of the optimal network maintenance problem employing our solution process to a fictitious network with 180 segments. In addition, Section 4 includes a comparison of common general purpose mathematical programming solvers against the proposed solution process with respect to effectiveness in reaching the optimum. Section 5 concludes.

2 Formulations of the maintenance problem

2.1 Binary programming formulation

Let the undirected network $G \equiv \{V, E, c_r, c_\delta, b_r, b_\delta\}$ describe the infrastructure network under study, where V is its node set, E is its arc set, and its parameters are given as follows:

$c_r : E \rightarrow \mathbb{R}^+$ is the local maintenance cost function.

$c_\delta : E \rightarrow \mathbb{R}^+$ is the excess cost of replacement function. To restore arc e to an “as new” condition the amount $c_r(e) + c_\delta(e)$ is required.

$b_r : E \rightarrow \mathbb{R}^+$ is the function assigning expected savings in emergency repair cost after local preventive maintenance.

$b_\delta : E \rightarrow \mathbb{R}^+$ is the function assigning additional expected savings after equipment renewal in excess to the ones resulting from local maintenance. If arc e is restored to an “as new” condition, then the total savings in expected emergency repair costs is given as $b_r(e) + b_\delta(e)$.

All costs and benefits are nonnegative.

The savings in emergency repair costs are usually calculated by multiplying the typically very high cost of emergency repair (includes possible liabilities) by the expected reduction in failure probability after maintenance. Risk assessment studies may provide estimates for failure probabilities without maintenance, after local maintenance, and after preventive replacement (renewal). Typically preventive

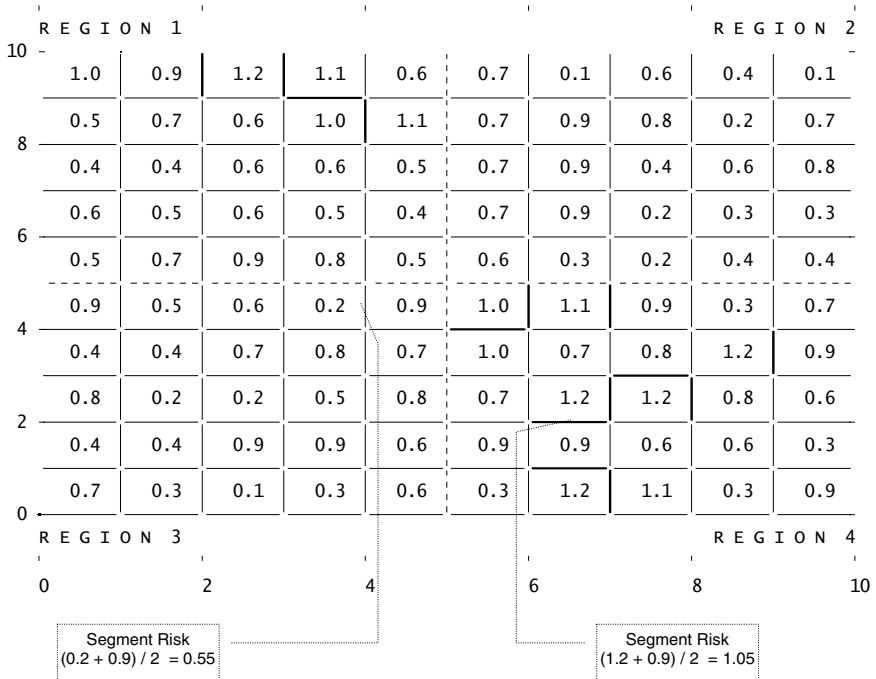


Fig. 2. Part of an infrastructure network divided in 4 regions. Each location is marked by its risk value with segment risk equal to the average of its adjacent locations. High risk arcs (thick lines) tend to form clusters.

replacement results in lower failure probabilities than local maintenance; therefore it is reasonable to assume b_δ takes nonnegative values only.

An example of an infrastructure network appears in Figure 2, where a fictitious natural gas pipeline a service area is depicted. High-risk segments, and thus candidates for preventive maintenance in the short term, are drawn in thick lines. It is clear that the ratio of high-risk segments to total number of segments should be small in a well-maintained network. The risk of each segment is calculated as the average of the risk in its surrounding neighborhoods. Hence, high-risk segments tend to be close to each other forming clusters, as in Figure 2. The network is divided into four regions within which setup discounts apply. This representation of the network may lead the reader to believe that another equity concern is also important to consider, transboundary risk. In many types of network maintenance problems transboundary risks are not significant. For instance in railroad safety, risk has more to do with the number of crossings than the concentration of traffic on the one or the other side of the rail line. In another example, natural gas distribution, risk is highly concentrated within a narrow strip of land surrounding the pipeline, thus the stakeholders are essentially located on the pipeline and not far away.

In some cases, however, transboundary risk is important. It is well known that divided highways tend to also divide cities in dissimilar sections with respect to income (or ethnic origin). Risks to property values from highway traffic may be very

pronounced on the one side of the network segment and non-existent on the other side. Policies guided by average risk considerations may lead to friction from both a rich area regarding noise reduction walls as inadequate and from its neighboring poor area regarding, say, noise reduction walls as expensive. In these transboundary risk cases, though, more policy options exist (e.g., raising a taller noise reduction wall only on the one side), thus a different analysis outside the scope of our paper is required.

NTD in maintenance investments are defined by the following two functions:

Contiguity discount function $d_c : 2^E \rightarrow \mathbb{R}^+$

$d_c(\{e, f\})$ is the cost that will not be replicated, if the contiguous arcs e and f are both replaced: $d_c(\{e, f\}) > 0 \Rightarrow \exists v \in V : e, f \in \theta(v)$, where the incidence sets $\theta(v)$ are defined as $\forall v \in V \theta(v) \equiv \{(v, h) : h \in V, (v, h) \in E\} \cup \{(h, v) : h \in V, (h, v) \in E\}$.

Two-way contiguity discounts are meant to exist for contiguous (or possibly near contiguous) arcs. No computational problem, however, arises in solving the maintenance problem, if the latter restriction is relaxed. Also, $\forall e \in E d_c(\{e\}) = 0$.

If two contiguous arcs e and f are to be replaced, then their replacement cost will be: $c_r(e) + c_\delta(e) + c_r(f) + c_\delta(f) - d_c(\{e, f\})$.

Setup cost discount function $d_s : E \times \Pi \rightarrow \mathbb{R}^+$

Where Π is a partition of physical network arcs to subregions according to the structure of setup costs: $\Pi \equiv \{R, R' \subset E : R \cap R' = \emptyset, \bigcup_{\Pi} R = E\}$. $d_s(e, R)$ is the savings in replacement costs for any arc $e \in R$ after the cost $c_s(R)$ is paid once. If $S \subset R$ are arcs to be replaced and the cost $c_s(R)$ is paid, then the resulting total cost of replacement excluding contiguity discounts for region R will be: $\sum_{e \in S} (c_r(e) + c_\delta(e) - d_s(e, R)) + c_s(R)$. The setup cost discounts are positive only in the neighborhood they refer to, i.e.: $e \notin R \Rightarrow d_s(e, R) = 0$.

It may be assumed that E is ordered according to some convention, so that $e_{(i)} \in E$ is the i^{th} section in the accepted order. Similarly, $R_{(j)} \in \Pi$ is the j^{th} region according to some accepted ordering. Define three cost vectors of magnitude $|E|$, $|E|$, and $|\Pi|$, respectively as: $(\underline{\mathbf{c}}_r)_i = c_r(e_{(i)})$, $(\underline{\mathbf{c}}_\delta)_i = c_\delta(e_{(i)})$, $(\underline{\mathbf{c}}_s)_j = c_s(R_{(j)})$. Also, \mathbf{B}_r , \mathbf{B}_δ , and \mathbf{D}_c are $[|E| \times |E|]$ matrices and \mathbf{D}_s a $[|E| \times |\Pi|]$ matrix given as follows:

$$(\mathbf{B}_r)_{ii} = b_r(e_{(i)}), \forall i \neq j (\mathbf{B}_r)_{ij} = 0; \quad (\mathbf{B}_\delta)_{ii} = b_\delta(e_{(i)}), \forall i \neq j (\mathbf{B}_\delta)_{ij} = 0$$

$$(\mathbf{D}_c)_{ij} = \begin{cases} \forall i < j & d_c(\{e_{(i)}, e_{(j)}\}) \\ 0 & \text{otherwise} \end{cases}; \quad (\mathbf{D}_s)_{ik} = d_s(e_{(i)}, R_{(k)})$$

Compacting cost-benefit data further with: $\underline{\mathbf{c}} = \begin{bmatrix} \underline{\mathbf{c}}_r \\ \underline{\mathbf{c}}_\delta \\ \underline{\mathbf{c}}_s \end{bmatrix} \geq \underline{\mathbf{0}}$, and $\mathbf{B} =$

$$\begin{bmatrix} \mathbf{B}_r & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_\delta & \mathbf{D}_c & \mathbf{D}_s \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \geq \mathbf{0} \text{ the total maintenance costs are given by:}$$

$$\mathbf{T} = \underline{\mathbf{c}}^T \underline{\mathbf{x}} = \underline{\mathbf{x}}^T \mathbf{B} \underline{\mathbf{x}} \quad (1)$$

Where the binary decision vector, $\underline{\mathbf{x}}$, has magnitude $(|E| + |E| + |II|)$ and is comprised of 0s and 1s as follows:

$\forall i : 1 \leq i \leq |E| \quad (\underline{\mathbf{x}})_i = 1$ if arc $e_{(i)}$ is to be locally maintained

$\forall i : |E| < i \leq 2 \cdot |E| \quad (\underline{\mathbf{x}})_i = 1$ if arc $e_{(i-|E|)}$ is to be replaced

$\forall i : 2 \cdot |E| < i \leq 2 \cdot |E| + |II| \quad (\underline{\mathbf{x}})_i = 1$ if the setup cost is paid in region $R_{(i-2 \cdot |E|)}$

The first row of blocks in the benefit matrix \mathbf{B} contains nonzero elements in the diagonal because no interaction in local-maintenance benefits is considered. In contrast, \mathbf{B} 's third row contains only zero elements, as the benefits of setup discounts are realized though interactions only. The second block row in \mathbf{B} contains no 1s in the diagonal and captures the structure of benefit interactions. Excess benefits in \mathbf{B}_d accrue only if both decision variables for local maintenance and excess work for renewal are set to 1.

Formulating \mathbf{D}_c as a lower triangular matrix assures no double counting of contiguity discounts and economizes in calculations. Finally, \mathbf{D}_s applies setup cost discounts to renewal jobs only and not to local maintenance, as appears to be the case in practice. No computational problem arises, however, if setup discounts can be realized in local maintenance.

Now the maintenance cost minimization problem becomes:

$$\text{(Prob. 1)} \quad \min_{\underline{\mathbf{x}} \in \{0,1\}^K} \{ \underline{\mathbf{c}}^T \underline{\mathbf{x}} - \underline{\mathbf{x}}^T \mathbf{B} \underline{\mathbf{x}} \} \text{ with } K = |E| + |E| + |II|$$

It is reasonable to assume that discounts do not exceed full cost in the following sense:

$$\underline{\mathbf{c}}_d > [\mathbf{D}_c \quad \mathbf{D}_s] \cdot \underline{\mathbf{1}} \quad (2)$$

Typically, contiguity and setup discounts are but a fraction of the total maintenance/reconstruction cost needed to return a defective segment to "as new" condition, if not an order of magnitude less. In addition, the cost of local maintenance is also a small fraction of reconstruction cost. Hence, in practice (2) is satisfied by a wide margin. Regularity assumption (2) is important though, in that it guarantees that no optimal solution to Problem 1 has the form: $\exists i : 1 \leq i \leq |E|, (\underline{\mathbf{x}}^*)_i < (\underline{\mathbf{x}}^*)_{i+|E|}$ which would have no meaning in practice.

The cost vector is likely to be easy to estimate. It is rather difficult, however, to estimate the Benefit Matrix, mainly because of uncertainties in the assessment of failure probabilities and consequently in the calculation of expected emergency repair costs. Handling uncertainty in benefits is discussed in Section 3.

2.2 Rhys-Balinski formulation

Equation (1) will be put in the form used by Balinski (1970) in his selection problem. Balinski (1970) generalized Rhys' (1970) optimization problem for investments with mutually exclusive costs but possibly not itemizable benefits. Without loss of generality we may use $H = \{1, 2, \dots, K\}$ as the set of possible maintenance investment opportunities in Problem 1. The cost of each investment $k \in H$ is $c'(k) = (\underline{\mathbf{c}})_k \geq 0$. The benefit from each investment set $\sigma \subset H$ is

$b'(\sigma) = \underline{\mathbf{x}}(\sigma)^T \mathbf{B} \underline{\mathbf{x}}(\sigma) \geq 0$, where $\underline{\mathbf{x}}(\sigma)$ is the indicator vector of set σ , i.e.:
 $\forall k \in \sigma \quad \left(\underline{\mathbf{x}}(\sigma) \right)_k = 1 \Leftrightarrow k \in \sigma$.

Note that function $b' : A \subset 2^H \rightarrow \mathfrak{R}^+$ is not difficult to construct. Only two-way benefit interactions are relevant to the NTD problem. Thus, the cardinality of the benefit function domain is not large: $|A| \leq \binom{|H|}{2}$. It is now clear that, with little computational effort, Problem 1 may be written as an optimization problem of itemizable costs and possibly non-itemizable benefits as:

$$\text{(Prob. 2)} \quad \sigma^* = \arg \max_{\sigma \in H} \left\{ b'(\sigma) - \sum_{k \in \sigma} c'(k) \right\}$$

2.3 Max flow – min cut formulation

Balinski (1970) has shown (see also Nemhauser and Wolsey, 1988, pp. 694–702, for a more instructive proof) that Problem 2 is equivalent to a maximum flow problem in an auxiliary directed network, heretofore called Cost Benefit Network (CBN). Let $\Omega \equiv \{N, A, c', b'\}$ be the CBN of Problem (1-2). It can be constructed as follows:

$N \equiv \{o, N_B \equiv \{\sigma \in A : b'(\sigma) > 0\}, N_C \equiv \{k \in H : c(k) > 0\}, q\}$, where o is a source node and q a sink node.

$A \equiv \{A_B \equiv \{(o, \sigma) : \sigma \in N_B\}, A_I \equiv \{(\sigma, k) : \sigma \in N_B, k \in \sigma \subset N_C\}, A_C \equiv \{(k, q) : k \in N_C\}\}$. With upper capacity limits as follows¹:

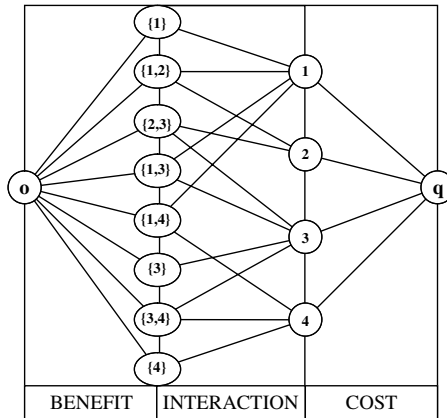


Fig. 3. Cost-benefit network example

¹ Typically, arcs in A_I have infinite capacity, but setting $C(e) = b'(\sigma) + \varepsilon$ for $e = (\sigma, k) \in A_I$ obviously does not change the nature of the problem as no arc in A_I would ever be part of the minimum cut, while it eliminates questions about the choice of a capacity that is high enough to be considered infinite.

$$\bar{C}(e) = \begin{cases} b'(\sigma) & \text{if } e = (o, \sigma) \in A_B \\ b'(\sigma) + \varepsilon & \text{if } e = (\sigma, k) \in A_I \\ c'(k) & \text{if } e = (k, q) \in A_C \end{cases} \quad \text{where } \varepsilon \in R^+ \text{ is a small quantity.}$$

An example of a CBN is depicted in Figure 3. Let a cut separating two distinct node sets in N be given as: $\forall X, Y \subset N; X \cap Y = \emptyset \quad \Gamma(X, Y) \equiv \{(i, j) \in A : i \in X, j \in Y\}$. Balinski (1970) has shown that the max-flow/min-cut problem below reveals the solution to Problem (1-2):

$$\text{(Prob. 3)} \quad X^* = \arg \min_{\{X \subset N : o \in X, q \notin X\}} \left\{ \sum_{e \in \Gamma(X, N-X)} \bar{C}(e) \right\}$$

In particular, the cost arcs in the minimum cut reveal the optimal solution to Problem 2, which is optimized by $\sigma^* = (N - X^*) \cap N_C$. Flow in the CBN corresponds to benefit and there is a one-to-one correspondence from network segment selection to flow in the A_C arcs. If the flow through an arc in A_C equals its capacity, it is called cost-limited and the net benefit of maintaining the corresponding segment is positive. Otherwise the flow is considered benefit-limited, thus corresponding segments are left out of the optimal maintenance list. Some arcs in A_C may appear to be cost-limited due to interactions with benefit-limited arcs. This problem is avoided if σ^* is calculated by employing a depth first search for arcs with positive interaction to cost-limited arcs.

It is well known that solving Problem 2 as a mixed integer program with a general purpose mathematical programming procedure, like branch and bound, will require exponentially increasing computation time. By contrast, referring for instance to Tarjans's (1986) analysis, general purpose network flow algorithms would require $O(|A|^2 \cdot \log(M))$ time for maximum capacity flow augmenting paths, with M being the maximum capacity limit for arcs in $A_B \cup A_C$. Alternatively, $O(|A|^2 \cdot |N|)$ time will be required for minimum cardinality flow augmenting paths.

Of the two approaches we recommend the second as it does not depend on the capacities taking integer values. Typically, rational capacities can be converted to integer after multiplication by a constant, this however increases M and is thus counterproductive. In the example we solve, the well known algorithm by Dinic (1970) is utilized. In the past 30 years there has been considerable progress in maximum flow algorithms. Goldberg (1998) in a recent overview of this progress cites algorithms that are faster by about one polynomial factor, closer to $O(|N|^2)$. These methods are clearly more scalable as they permit the timely solution of much bigger problems. In our implementation, we used a free version of Dinic's Algorithm, available from the Center for Discrete Mathematics and Theoretical Computer Science at Rutgers University (dimacs.rutgers.edu), which dominated other approaches taking into account convenience and code reliability.

It is important to note that in a well maintained network the majority of segments will not be even considered for maintenance. An obvious simplification of the problem we consider arises from the a priori elimination of segments in excellent condition. For these segments the risk of failure is so small, that even if all possible network dependencies associated with them are counted the net benefit of their replacement is still negative. Depending on the state of the network and not only

on its size the set of remaining segments (ambiguous with respect to inclusion to the optimal maintenance list) can then become the input to the proposed selection procedure. Hence, even though the total number of segments in a utility network may be in the order of hundreds of thousands, one expects that the segments for which inclusion to the optimal maintenance list is unclear would be an order of magnitude less. In the discussion of our illustrative example we show that results for sizable networks can be obtained with an off-the-self computer in minutes. We expect that various factors affect computational time, including the state of the network, the size of network dependencies, and the proximity of poor-condition segments. Investigating in depth the determinants of computational requirements, when our procedure is applied to a real utility network, is left outside the scope of this paper.

CBN optimization leads to exact solutions. Enumeration techniques, like branch and bound, and more often crude heuristics terminate with approximations of the optimal solution. It is important to note, that the values of total maintenance costs for near-optimal solutions may lie close to the exact optimum, but the near-optimal arc list may be quite distinct from the optimal list in a geographic sense. Obtaining the true optimal solution is an important advantage of the proposed procedure, when it is applied to geographic-distribution-sensitive risk management problems. Furthermore, exact solutions reduce significantly the latitude for manipulation in maintenance prioritization, as other parties using the same parameters may reproduce the unique results of the optimization procedure.

2.4 Enforced maintenance formulation

Finally, we show that requiring maintenance of certain segments, possibly due to equity considerations, is very easy to do within the solution framework. Let \underline{e} be the vector of enforced maintenance, then the solution space of Problem 1 is transformed:

$$\text{(Prob. 1.a)} \quad \min_{\underline{x} \geq \underline{e}} \{ \underline{c}^T \underline{x} - \underline{x}^T \mathbf{B} \underline{x} \}$$

But now \underline{x} can be written as:

$$\underline{x} = \underline{x} - \underline{e} + \underline{e} = \underline{y} + \underline{e} \quad \text{where} \quad \underline{0} \leq \underline{y} \leq \underline{1} \tag{3}$$

And the objective function of Problem 1.a becomes:

$$\underline{c}^T (\underline{y} + \underline{e}) - (\underline{y} + \underline{e})^T \mathbf{B} (\underline{y} + \underline{e}) = (\underline{c}^T - \underline{e}^T \mathbf{B}^T - \underline{e}^T \mathbf{B}) \underline{y} - \underline{y}^T \mathbf{B} \underline{y} + (\underline{c}^T \underline{e} - \underline{e}^T \mathbf{B} \underline{e}) \tag{4}$$

Note that the third factor in the right part of (4) is a constant, hence the objective of Problem 1.a has the same form as the objective in Problem 1. Consequently, enforcing maintenance on some arcs does not add difficult to deal with complications to the proposed solution process.²

² Note that significant complications arise when there are negative externalities in the benefit function, in addition to the positive externalities assumed here. For maintenance problems, one would normally expect only positive externalities. In the case of simultaneous negative and positive externalities, solving the maintenance problem would probably require an altogether different methodology (e.g., global optimization).

3 Sensitivity analysis

Brumelle and Granot (1993) have analyzed in detail the effect of parametric changes in the solution set of a “repair-kit” problem. Problem 2 belongs to the family of “repair-kit” problems they analyze and their results are applicable to sensitivity analysis of the problem studied here. In the next we focus first on the effect of a uniform increase in maintenance benefits in response to a reevaluation of expected emergency repair costs. Then we discuss the selection of total maintenance budget by varying cost of capital. We show, also, that in both cases parametric variation leads to nested solutions.

There exist many examples of parameters that affect all service region benefits in the same direction: for instance: upwards revision of the probability of failure after an unknown failure mode is discovered, increases in emergency repair labor costs, increases in possible liabilities for section failures. The revision of other parameters may or may not produce a uniform upwards revision of benefits. For instance, changes in economic value and population densities do not always occur in a uniform way. Some city neighborhoods experience positive growth, but other neighborhoods may experience zero growth or decline. Uniform revisions of benefits, however, are very common in infrastructure maintenance problems and their effect is very conveniently obtained by the CBN procedure. This property of our solution framework is shown in the following.

First, submodularity is verified by showing the following:

Proposition 1. If $\underline{x} \in \{0, 1\}^n$, $\underline{d} \in \mathbb{R}^n$, and $\mathbf{Q} \geq \mathbf{0}$ is a nonnegative $[K \times K]$ matrix, then $F(\underline{x}; \mathbf{Q}, \underline{d}) = \underline{d}^T \underline{x} - \underline{x}^T \mathbf{Q} \underline{x}$ is submodular (i.e.: $F(\underline{x}; \mathbf{Q}, \underline{d}) + F(\underline{x}'; \mathbf{Q}, \underline{d}) \geq F(\underline{x} \vee \underline{x}'; \mathbf{Q}, \underline{d}) + F(\underline{x} \wedge \underline{x}'; \mathbf{Q}, \underline{d})$).

Where $\underline{x} \wedge \underline{x}'$, $\underline{x} \vee \underline{x}'$ symbolize respectively greatest lower and least upper bounds, following Topkis' (1978) notation. Topkis (1978) provides also descriptions of submodular function properties and examples.

Proof. Without loss of generality we put vectors \underline{x} , \underline{x}' in the following form:

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \\ \underline{x}_4 \end{bmatrix}, \underline{x}' = \begin{bmatrix} \underline{x}'_1 \\ \underline{x}'_2 \\ \underline{x}'_3 \\ \underline{x}'_4 \end{bmatrix} \quad \text{so that} \quad \begin{cases} \underline{x}_1 = \underline{x}'_1 = \underline{1} \\ \underline{x}_2 = \underline{1}, \underline{x}'_2 = \underline{0} \\ \underline{x}_3 = \underline{0}, \underline{x}'_3 = \underline{1} \\ \underline{x}_4 = \underline{x}'_4 = \underline{0} \end{cases}$$

$$\text{and } \underline{x} \wedge \underline{x}' = \begin{bmatrix} \underline{1} \\ \underline{0} \\ \underline{0} \\ \underline{0} \end{bmatrix}, \underline{x} \vee \underline{x}' = \begin{bmatrix} \underline{1} \\ \underline{1} \\ \underline{1} \\ \underline{0} \end{bmatrix} \quad (5)$$

where $\underline{0}$ and $\underline{1}$ are zero and one subvectors. Matrix \mathbf{Q} is put accordingly in the following form:

$$\mathbf{Q} = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} & Q_{14} \\ Q_{21} & Q_{22} & Q_{23} & Q_{24} \\ Q_{31} & Q_{32} & Q_{33} & Q_{34} \\ Q_{41} & Q_{42} & Q_{43} & Q_{44} \end{bmatrix}$$

Now by regular algebra:

$$\begin{aligned}
 F(\underline{\mathbf{x}}; \mathbf{Q}, \underline{\mathbf{d}}) + f(\underline{\mathbf{x}}'; \mathbf{Q}, \underline{\mathbf{d}}) &= F(\underline{\mathbf{x}} \wedge \underline{\mathbf{x}}'; \mathbf{Q}, \underline{\mathbf{d}}) + F(\underline{\mathbf{x}} \vee \underline{\mathbf{x}}'; \mathbf{Q}, \underline{\mathbf{d}}) \\
 &+ \begin{array}{c} \left| \begin{array}{c} \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ \underline{\mathbf{0}} \end{array} \right|^T \\ \left| \begin{array}{cccc} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{32} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{32} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array} \right| \left| \begin{array}{c} \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ \underline{\mathbf{0}} \end{array} \right| \end{array} \quad (6)
 \end{aligned}$$

□

Therefore, the objective function of Problem 1 is submodular. This well known result is offered to improve our exposition. Alternatively, one could verify the conditions of Topkis's (1978) Lemma 3.1.

Proposition 2. The net cost function of Equation (1) has the antitone differences property:

$$\begin{aligned}
 \underline{\mathbf{x}} \leq \underline{\mathbf{y}}, \mathbf{B} \leq \mathbf{B}' &\Rightarrow (\underline{\mathbf{c}}^T \underline{\mathbf{y}} - \underline{\mathbf{y}}^T \mathbf{B} \underline{\mathbf{y}}) - (\underline{\mathbf{c}}^T \underline{\mathbf{x}} - \underline{\mathbf{x}}^T \mathbf{B} \underline{\mathbf{x}}) \\
 &\geq (\underline{\mathbf{c}}^T \underline{\mathbf{y}} - \underline{\mathbf{y}}^T \mathbf{B}' \underline{\mathbf{y}}) - (\underline{\mathbf{c}}^T \underline{\mathbf{x}} - \underline{\mathbf{x}}^T \mathbf{B}' \underline{\mathbf{x}}) \quad (7)
 \end{aligned}$$

Proof. It suffices to show:

$$\underline{\mathbf{x}}^T \mathbf{B} \underline{\mathbf{x}} - \underline{\mathbf{y}}^T \mathbf{B} \underline{\mathbf{y}} \geq \underline{\mathbf{x}}^T \mathbf{B}' \underline{\mathbf{x}} - \underline{\mathbf{y}}^T \mathbf{B}' \underline{\mathbf{y}} \Leftrightarrow \underline{\mathbf{y}}^T (\mathbf{B}' - \mathbf{B}) \underline{\mathbf{y}} - \underline{\mathbf{x}}^T (\mathbf{B}' - \mathbf{B}) \underline{\mathbf{x}} \geq 0 \quad (8)$$

Breaking down the vectors $\underline{\mathbf{x}}, \underline{\mathbf{y}}$ and the matrix $\mathbf{B}' - \mathbf{B} = \Delta \mathbf{B}$ into their components, and noting that $\forall i, j$ $(\underline{\mathbf{x}})_i \leq (\underline{\mathbf{y}})_i$, $\Delta \mathbf{B}_{ij} \geq 0$ one obtains:

$$\underline{\mathbf{y}}^T (\Delta \mathbf{B}) \underline{\mathbf{y}} - \underline{\mathbf{x}}^T (\Delta \mathbf{B}) \underline{\mathbf{x}} = \sum_{i,j} \Delta \mathbf{B}_{ij} \left((\underline{\mathbf{y}})_i (\underline{\mathbf{y}})_j - (\underline{\mathbf{x}})_i (\underline{\mathbf{x}})_j \right) \geq 0 \quad \square$$

Let us now consider the impact of changes in benefit arc capacities on the optimal solution. It may be intuitive that nonnegative changes in benefit arc capacities only of the CBN will result in a new minimum cut that contains all cost arcs obtained using the original benefit parameters. This important fact was established by Topkis (1978), and follows directly from his results as we now note.

The parameter set of Problem 1 is $\{\underline{\mathbf{c}} \geq \underline{\mathbf{0}}, \mathbf{B} \geq \underline{\mathbf{0}}\}$. Consider the effect of the following parametric change:

$$\underline{\mathbf{y}}' = \arg \min_{\{0,1\}} \{\underline{\mathbf{c}}^T \underline{\mathbf{y}} - \underline{\mathbf{y}}^T \mathbf{B}' \underline{\mathbf{y}}\} \quad (9)$$

Proposition 3. $\mathbf{B}' \geq \mathbf{B} \Rightarrow \underline{\mathbf{y}}^* \geq \underline{\mathbf{x}}^*$

Proof. Note that $\{0,1\}^K$, the space the decision vectors are constrained in, is a lattice. In addition, the parameter $\mathbf{B} \in \mathfrak{R}^{K \cdot K}$ belongs to a partially ordered set. Thus, according to Theorem 6.1 in Topkis (1978), after the parametric change in (9), the optimal vector for the original parameter set is no greater than the new optimal vector, provided that the following two conditions hold. First, the objective function is submodular, which is true by Proposition 1. And second, the antitone differences property holds, which is true by Proposition 2. □

Thus, if \mathbf{B} is the most conservative benefit estimate and \mathbf{B}' is an upward revision of \mathbf{B} , then all work prescribed by $\underline{\mathbf{x}}^*$ will be prescribed also by maintenance plan $\underline{\mathbf{y}}^*$.

\underline{y}^* may be computed faster if \underline{x}^* is known by computing $\underline{y}^* - \underline{x}^*$ only. This process may be extended so that the optimal maintenance plan is computed in a stepwise fashion. That is start with a very conservative estimate of benefits, revise them upwards in steps, and in the process increment the optimal maintenance plan. Brumelle and Granot (1993) elaborate on this approach.

The total value, V , of risk reduction (typically the net present value of expected liability reduction) and the associated annual budget, W , are given as:

$$V(\underline{x}) = \underline{x}^T \begin{bmatrix} \mathbf{B}_r & \mathbf{0} & \mathbf{0} \\ \mathbf{B}_d & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \underline{x} \quad (10)$$

$$W(\underline{x}) = \underline{c}^T \underline{x} - \underline{x}^T \mathbf{D} \underline{x} = \underline{x}^T (\underline{c} - \mathbf{D} \underline{x}) \quad (11)$$

Of particular interest is finding a budget limit \overline{W} , which is not exceeded by the total cost of maintenance under a maintenance plan \underline{x} .

$$W(\underline{x}) = \underline{c}^T \underline{x} - \underline{x}^T \mathbf{D} \underline{x} = \overline{W} \quad (12)$$

Fortunately, Problem 1 need not be solved under constraint (12), but its Lagrangian relaxation defines the solution to the budgeting problem. In fact, if we formulate the Lagrangian relaxation of (12) in Problem 1, we obtain:

$$(\text{Prob. 4}) \quad \underline{x}^* = \arg \min \{ \underline{c}^T \underline{x} - \underline{x}^T \mathbf{B} \underline{x} + \lambda (W(\underline{x}) - \overline{W}) \} \text{ where } \lambda \in R^+$$

We can interpret λ as the network operator's cost of capital. Varying λ from zero to some upper bound gives rise to a feasible solution $\underline{x}^*(\lambda)$ and an optimal budget $\overline{W}(\underline{x}^*(\lambda))$ for each value λ (the constant \overline{W} in Prob. 4 has no effect on the solution value, of course). This approach yields the efficient frontier in a natural way to increasing cost of capital and avoids problems of budgeting inefficiency arising from out of frontier solutions and related duality gaps associated with Problem 1. From a policy perspective, it is the preferred approach and the one we follow. Reformulating Problem 4 by multiplying with $\gamma = 1/(1 + \lambda)$, the results of Proposition 3 become very clear in terms of the cost of capital approach embodied in Prob. 4.

$$(\text{Prob. 5}) \quad \underline{x}_\gamma^* = \arg \min \{ \underline{c}^T \underline{x} - \underline{x}^T (\mathbf{D} + \gamma(\mathbf{B} - \mathbf{D})) \underline{x} \} \text{ where } \gamma \in (0, 1]$$

Proposition 4. $\gamma < \gamma' \Rightarrow \underline{x}_\gamma^* \leq \underline{x}_{\gamma'}^*$ and $W(\underline{x}_\gamma^*) \leq W(\underline{x}_{\gamma'}^*)$

Proof. The first part follows immediately from Proposition 3. For the second we use the optimality of \underline{x}_γ^* to show that:

$$\begin{aligned} W(\underline{x}_{\gamma'}^*) - \gamma \underline{x}_{\gamma'}^{*T} (\mathbf{B} - \mathbf{D}) \underline{x}_{\gamma'}^* &= \underline{c}^T \underline{x}_{\gamma'}^* - \underline{x}_{\gamma'}^{*T} (\mathbf{D} + \gamma(\mathbf{B} - \mathbf{D})) \underline{x}_{\gamma'}^* \\ &\geq W(\underline{x}_\gamma^*) - \gamma \underline{x}_\gamma^{*T} (\mathbf{B} - \mathbf{D}) \underline{x}_\gamma^* \Rightarrow \\ W(\underline{x}_{\gamma'}^*) - W(\underline{x}_\gamma^*) &\geq \gamma [\underline{x}_{\gamma'}^{*T} (\mathbf{B} - \mathbf{D}) \underline{x}_{\gamma'}^* - \underline{x}_\gamma^{*T} (\mathbf{B} - \mathbf{D}) \underline{x}_\gamma^*] \geq 0 \end{aligned} \quad (13)$$

The last part of (13) follows from the fact that $\mathbf{B} - \mathbf{D} > \mathbf{0}$ and from the first part of Proposition 4. \square

Hence, solving for successively higher γ , the optimal management of the network requires higher budgets also. It is a very convenient property of our solution process that as increasing values of γ (decreasing for the cost of capital) are tried

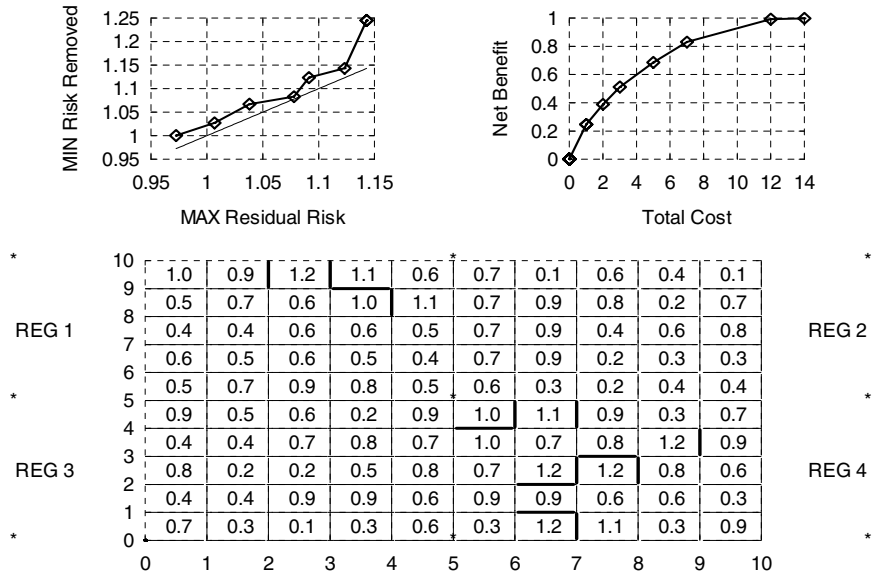


Fig. 4. Overview of solution in the absence of NTD (replacement cost = 1, contiguity discount = 0, setup discount = 0). Top panels show the effect of choosing successively lower benefit-cost ratios. Thick lines denote segments in the optimal maintenance plan

out the optimal maintenance list is expanded and is not altered radically. The latter property is often referred to as the “Nestedness Property.” Furthermore, due to cost structure submodularity, as γ values increase one conveniently obtains the cost-benefit frontier with few calculations.

4 Illustrative example

The example described in Figure 1 and in Figure 2 is worked on using the proposed method that prescribes first to transform Problem 1 to a maximum flow problem and then to solve using an established algorithm. Particular attention is paid to the tradeoffs in efficiency and equity as the benefit-cost ratio varies.

Figure 4 depicts the geographic distribution of the optimal maintenance list when NTD are not taken into account and the cost of capital is zero (non-binding budget constraint). 14 out of 180 (7.8%) segments are in the maintenance list. From the 14 segments in the maintenance list, 11 are contiguous to another maintenance list arc (78.6%). Having so many contiguous arcs despite the absence of NTD can be explained by the spatial dependencies in the risk structure. The reader is reminded that the risk of each segment is the average risk of its surrounding neighborhoods. Even though neighborhood risk was simulated as geographically independent, segment risk by construction has an obvious dependence on geographic proximity. As high risk segments are contiguous, segments in the optimal maintenance list are also contiguous. This geographic dependence in segment risk is often a very

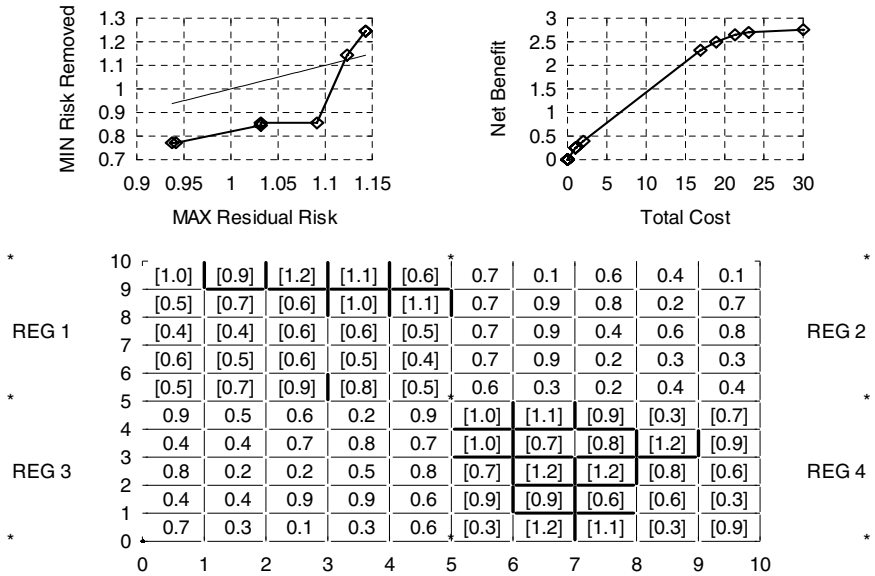


Fig. 5. Overview of solution when maintenance cost structure includes NTD (replacement cost = 1, contiguity discount = 0.05, setup discount = 0.15). Top panels show the effect of choosing successively lower benefit-cost ratios. Thick lines denote segments in the optimal maintenance plan. The brackets in regions 1 and 4 denote that setup discounts apply

realistic assumption. When NTD are added to the maintenance cost structure the pattern of contiguous maintenance list segments will be more pronounced.

In addition, Figure 4 depicts two panels with the equity and efficiency tradeoffs as the cost of capital increases and in turn total budget decreases. In this “knapsack” problem, optimal maintenance lists obtained by varying λ lie on the benefit cost efficiency frontier. A property that will also apply after NTD are added. The equity tradeoff is studied using a simple metric, the difference between the minimum risk planned for elimination (using the maintenance budget) and the maximum risk of segments outside the maintenance list. The minimax points lie on an increasing curve situated above the identity line (thin line in top left panel of Fig. 4). Thus, the minimum eliminated risk is higher than the maximum residual risk. This property makes it easier for the network operator to explicate its maintenance policy.

In the same format as in Figure 4, Figure 5 depicts the results one obtains after NTD are taken into account. 34 out of 180 segments are now scheduled for maintenance. In addition, setups are prescribed for regions 1 and 4, where neighborhood risk is put in brackets. Now all but one segments in the maintenance list are contiguous to another maintenance list segment. With respect to efficiency, tradeoffs are clear as all solutions generated by varying λ belong to the benefit-cost efficiency frontier. Figure 6 verifies the “Nestedness Property”, of the optimal maintenance list as cost of capital increases and total budget decreases. Note, however, that risk rank does not define entry order to the optimal solution as λ increases.

Benefit Cost Ratio vs Entry Order in Solution

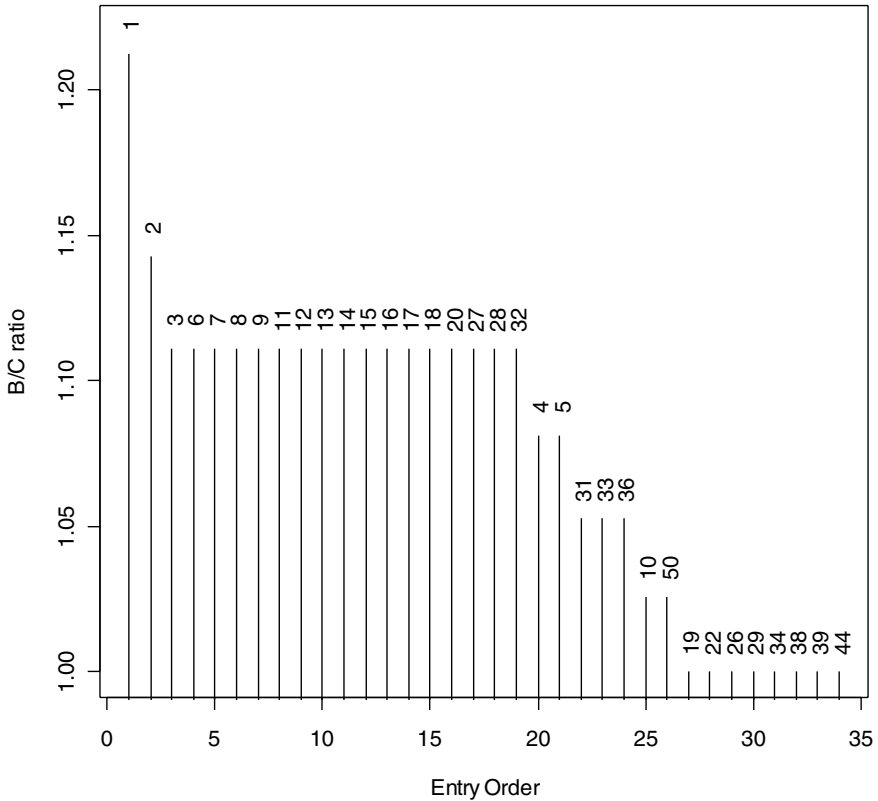


Fig. 6. Successive optimal set augmentation as benefit / cost ratio decreases (nestedness property). Labels denote segment risk rank (e.g., 1 is the highest risk segment). All segments with equal B/C ratio enter the solution together

With respect to equity, note in Figure 5 that the minimax risk points form a curve that is still increasing, but now crosses the identity line. Lower risk segments are scheduled for maintenance, due to their being contiguous to other segments or on the regions where special setups are prescribed. At the same time, higher risk segments that are geographically isolated are left unmaintained. This might create friction with residents of neighborhoods exposed to the higher residual risks. Equity tradeoffs are clear in the decision process we propose, so network operator managers may take them into account. Consider two options to influence residual risk distribution. First, enforce maintenance of segments with very high risk independent of whether or not they enter the optimal (most efficient) maintenance list. And second, choose points on the minimax curve that have low disparities between eliminated and non-eliminated risks. For instance, choose the point (Min Removed: 0.85, Max Residual, 1.03) over the point (Min Removed: 0.85, Max Residual, 1.09). Alternatively, network operator managers may opt for making clear the shared benefits from the increased efficiencies due to NTD.

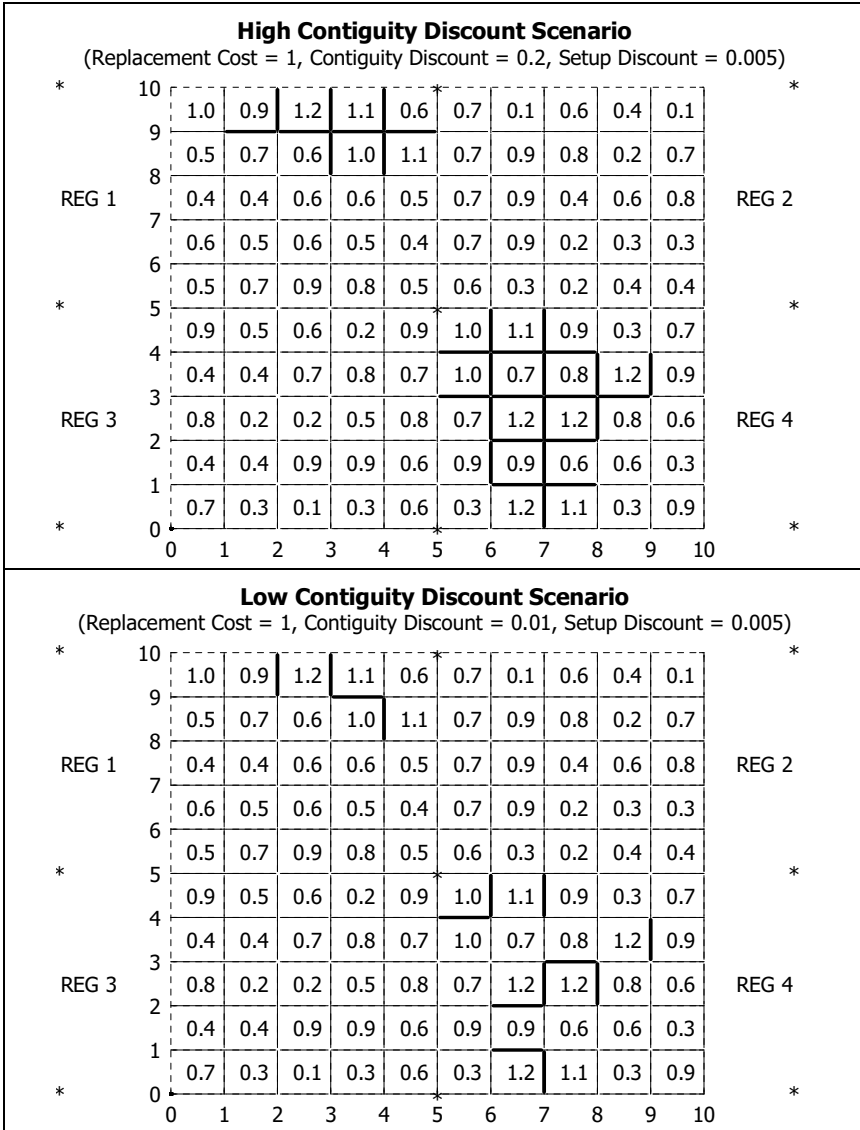


Fig. 7. Effect of varying contiguity discount magnitude. Bold segments to be maintained

Experiments were conducted with different cost parameters for discount and contiguity. Figure 7 shows the effect of going to high (top) and low (bottom) contiguity discounts magnitude. High contiguity discount leads to the optimal maintenance list forming two big clusters of segments. Regarding equity, we found in this scenario high disparity between eliminated risk (min: 0.71) and residual risk (max: 0.93). When contiguity discounts are lower the segments in the maintenance list are not as well connected. With respect to equity, disparities exist but are not as pronounced. Figure 8 shows the effect of going to high (top) and low (bottom)

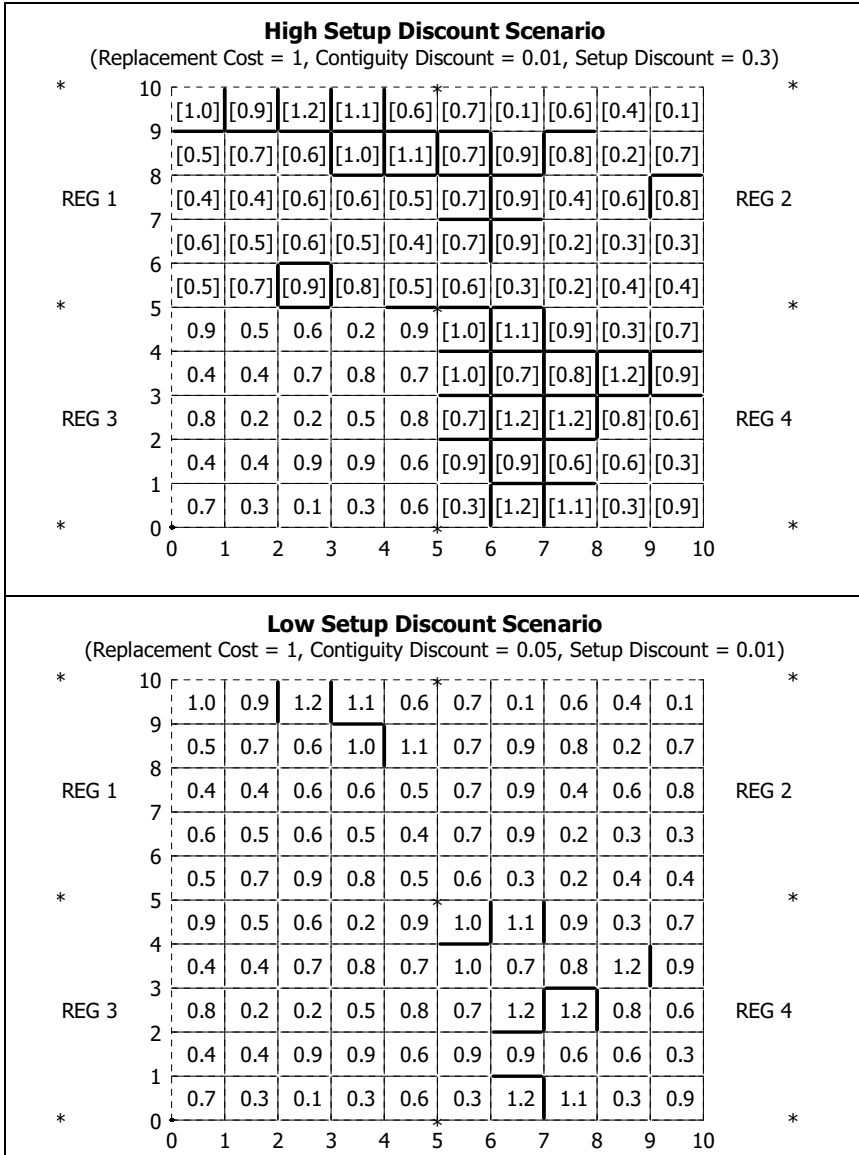


Fig. 8. Effect of varying setup discount magnitude. Bold segments to be maintained

setup discount magnitudes. High setup discounts lead to an optimal maintenance list prescribing setups in 3 out of 4 regions. With respect to equity, disparities are also very high (Min Eliminated: 0.68, Max Not Eliminated, 0.93).

All the scenarios above were tested in a small network of 180 segments and 4 regions, which generated a CBN of about 500 nodes and 1,200 arcs. Finding the Pareto frontier using a conventional Linux server and unoptimized code was achieved in less than a second. The Pareto frontier of a larger version of this network

Table 1. Comparison of solutions obtained by the proposed method to solutions obtained using common general purpose solvers. Maximum net benefit is rarely approached by general purpose solvers

NTD scenario	Proposed method	Minos	Loqo	Donlp2
Base	2.7550	2.6970 (−2.1%)	2.7550 (−0.0%)	2.6970 (−2.1%)
High contiguity	3.8340	3.7309 (−2.7%)	3.7869 (−1.2%)	3.7868 (−1.2%)
High setup	8.3690	8.0892 (−3.3%)	8.3096 (−0.7%)	8.0892 (−3.3%)

(of similar size to what may be encountered in practical applications) with roughly 6,000 segments and about 18,000 CBN nodes was calculated in less than 10 minutes on the same conventional machine.

Finally, we compared our solutions to the solutions obtained by commonly used general purpose mathematical programming solvers. These solvers assume convexity and search for a local optimum which often is an interior point. In Problem 1, one searches for the minimum of a concave function which lies always on the boundaries. This property of Problem 1 poses testing challenges to general purpose solvers as can be seen in Table 1. The best performance was obtained by the solver *Loqo* in all three scenarios considered, but in most cases general purpose solvers failed to find the optimal solution which is always obtained by the proposed procedure.

5 Conclusions

Network topology dependencies apply to physical asset maintenance of the network industries and have a wide range of applications, including natural gas, oil, and water pipelines as well as railroad, highway, underground cable and sewer networks. Changes in the regulatory environment of network industries are expected to increase incentives for efficiency in maintenance and therefore drive decision makers away from crude heuristics and towards methodologically sound decision processes. The problem of optimal infrastructure network maintenance under network topology dependencies is formulated and solved by a procedure that is computationally efficient, facilitates sensitivity analysis, and avoids manipulability.

The latter attributes of the proposed maintenance planning process are essential for success in implementing any prioritization procedure for infrastructure maintenance. Firstly, infrastructure networks are large in size and comprised by a vast number of elements. Therefore, computational efficiency is a requirement. Secondly, risk estimates for the failure of network sections are rarely robust to problem parameters and are typically given within ranges of uncertainty. Sensitivity analysis is thus a necessary element of maintenance prioritization. Under the proposed solution, the most frequently occurring parametric uncertainties may be analyzed

very efficiently. Finally, prioritization of physical risks that may involve loss of life, serious injuries, and significant economic disruption is a decision likely to be challenged by many and diverse parties. The proposed selection process produces exact solutions, which may be replicated by independent analysts working with the same problem parameters. Near optimal solutions may be close in the objective function range, but far apart in a geographic sense. Thus heuristic procedures, not producing the same solution every time they are applied, may be seen as arbitrary and manipulable. By contrast, our solution framework generates nested solutions that do not vary radically after parametric changes. Thus, our approach provides both a logical structure and intuitively appealing and explicable results.

Our work stresses the contribution of Operations Research and Management Science not only in developing quantitative analysis procedures for decision support, but also in integrating decision support systems with organizational and interorganizational decision processes. This integration is not a trivial matter and requires in-depth knowledge of the problem context and of quantitative analysis. An essential part of this integration is managing complexity and transparency. In our paper we permit solution process complexity when transparency is not a requirement in order to achieve computational efficiency. By contrast, when transparency is required, for instance when risk equity is questioned, tradeoffs are made as clear as possible using sensitivity analysis.

Optimization procedures are designed to solve problems. They may, however, generate problems in their implementation, raising questions of manipulability, for example, if applied inappropriately. These side effects of quantitative decision processes are not unpredictable and are often avoidable. Considerable effort has been allocated so that the solution procedure fits the risk management problem in hand. It is often difficult to legitimate physical risk allocation to a multitude of affected stakeholders. For network maintenance problems involving risk to the public, it is important that procedures used be able to accommodate and transparently display the results of budget restrictions, equity considerations and appropriate cost factors. The procedure described here appears to satisfy these requirements.

References

- Balinski ML (1970) On a selection problem. *Management Science* 17(3): 230–231
- Brumelle S, Granot D (1993) The repair kit problem revisited. *Operations Research* 41(5): 994–1006
- Crew M, Kleindorfer P (2002) Regulatory economics: twenty years of progress? *Journal of Regulatory Economics* 21(1): 5–22
- Dinic EA (1970) Algorithm for solution of a problem of maximum flow with power estimation. *Soviet Math Dokl* 11: 1277–1280
- Goldberg AV (1998) Recent developments in maximum flow algorithms. Technical Report #98-045, NEC Research Institute, Princeton, NJ
- Lawler E (2001) *Combinatorial optimization: networks and matroids*. Dover (Reprint of Holt, Rinehart & Winston, 1976 edn), Mineola, NY
- Mamer JW, Smith SA (1982) Optimizing field repair kits based on job completion rate. *Management Science* 28(11): 1328–1333

- Nemhauser GL, Wolsey LA (1988) Integer and combinatorial optimization. Wiley, New York
- Rhys JK (1970) A selection problem of shared fixed costs and network flows. *Management Science* 17(3): 200–207
- Tarjan RE (1986) Algorithms for maximum network flow. *Mathematical Programming Study* 26: 1–11
- Topkis DM (1978) Minimizing a submodular function on a lattice. *Operations Research* 6(2): 305–321