2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

On the Exact Solution of a Class of Stackelberg Games

Alexis L. Motto, *Member, IEEE*

***Abstract*— This paper presents a mixed-integer linear reformulation for a subclass of two-level nonlinear discrete-continuous decision problems also known as Stackelberg strategic games. Solving the original (two-level and nonlinear) model is algorithmically challenging. In fact, even global solvers for standard (or single-level) mixed-integer nonlinear programs are difficult to implement and relatively inaccessible. The proposed transformation allows solving this class of Stackelberg games using mixed-integer linear programming theory for which optimized solvers with large-scale computational capabilities have been reported. Furthermore, the proposed reformulation is achieved without adding any new integer variables to the original model, a desired computational property in case a large portion of the search tree has to be enumerated.**

I. INTRODUCTION

Consider the following class of optimization problems, also known as Stackelberg games or bilevel programs [1]:

$$
\begin{aligned}
[\mathcal{P}_1] \quad & \min_{u} \; c_0^T u_1 + c_1^T \widehat{x}_1(u_2) + c_2^T(u_2)\widehat{x}_2(u_2) && \text{(1a)}\\
& \text{s.t.} \; h(u, \widehat{x}(u_2), \lambda) \geq 0 && \text{(1b)}\\
& u \in \mathcal{U}_1 \times \mathcal{U}_2 && \text{(1c)}\\
& \text{where} \; \widehat{x}(u_2) \in \arg\min_{x} \; d_1^T x_1 + d_2^T(u_2)x_2 && \text{(1d)}\\
& \text{s.t.} \; A_1 x_1 + A_2(u_2)x_2 = b(u_2) && \text{(1e)}\\
& x_1, x_2 \geq 0 && \text{(1f)}
\end{aligned}
$$

where n_1, n_2, m, p_{11}, p_{12}, and p_{12} are nonnegative integers, $p_1 = p_{11}+p_{12}$, $x_1 \in \mathbb{R}^{n_1}$, $x_2 \in \mathbb{R}^{n_2}$, $u_1 \in \mathcal{U}_1 \subset \{0,1\}^{p_{11}} \times \mathbb{R}^{p_{12}}$, $u_2 \in \{0,1\}^{p_2}$, $b(u) \in \mathbb{R}^m$, $c_1 \in \mathbb{R}^{n_1}$, $c_2(u) \in \mathbb{R}^{n_2}$, $d_1 \in \mathbb{R}^{n_1}$, $d_2 \in \mathbb{R}^{n_2}$, $A_1 \in \mathbb{R}^{m\times n_1}$, $A_2 \in \mathbb{R}^{m\times n_2}$, h, b, c_2, d_2 and A_2 are affine in their respective arguments.

Program (1) models a class of decision situations involving two players who try to optimize their respective objective functions over a jointly dependent set. The first player (or the leader or the upper-level decision maker) has explicit control only on a subset of the decision variables, denoted by u, and the second player (or the follower or the lower-level decision maker) has explicit control on the complement subset of the decision variables, denoted by x. Assuming that both players are rational, the first player must take into account the optimal reaction vector of the second player when selecting his decision vector. Furthermore, note that for a fixed leader's decision sub-vector u_2, the lower-level problem is a linear program.

This work was supported by a Postdoctoral Fellowship Award of the Natural Sciences and Engineering Research Council (NSERC), Canada.

The author is with the Department of Biomedical Engineering, McGill University, 3775 University Street, Montreal, QC, Canada H3A 2B4: alexis.motto@mcgill.ca.

Remark 1: For conciseness, we will hereafter denote by $[\mathcal{P}_1]$ the class of mathematical programs including (1) and any mathematical program transformable to (1) by linearization, using possibly additional binary variables, continuous variables, or constraints.

The following assumptions are supposed to hold throughout this paper.

Assumption 1: For simplicity or practical purposes, we assume that the convex hulls of the sets defined by (1c), and (1e)–(1f), and by (1b), (1c), and (1e)–(1f) are nonempty and bounded, which ensures that (1) is well posed.

Assumption 2: The objective functions z_1 and z_2 are bounded for all u and x satisfying Assumption 1.

Assumption 3: For any given upper-level decision subvector $u_2 \in \mathcal{U}$, the optimal reaction set of the lower-level problem is bounded.

Assumption 4: Given any upper-level decision vector $u_0 \in \mathcal{U}$, the linear independence constraints qualification (LICQ) holds for (1d)–(1f).

An example of application of the aforementioned optimization model arises from the need to maximize some measure of network security under intentionally disruptive threat [2], [3].

The main objective of this paper is to cast $[\mathcal{P}_1]$, which is a two-level mathematical program, into an equivalent one-level mixed integer linear program for which efficient global solvers exist.

Next, we present an exact solution approach to (1), which requires no additional binary variables.

II. MAIN RESULTS

The main contributions of this paper are summarized in the following results.

Theorem 1: Problem (1) is equivalent to the following mathematical program in standard (i.e. one-level) format:

$$
\begin{aligned}
[\mathcal{P}_2] \quad & \min_{u,x,\lambda} \; c_0^T u_1 + c_1^T x_1 + c_2^T(u_2)x_2 && \text{(2a)}\\
& \text{s.t.} \; h(u, x, \lambda) \geq 0 && \text{(2b)}\\
& u \in \mathcal{U}_1 \times \mathcal{U}_2 && \text{(2c)}\\
& A_1 x_1 + A_2(u_2)x_2 = b(u_2) && \text{(2d)}\\
& x_1, x_2 \geq 0 && \text{(2e)}\\
& \lambda^T A_1 \leq d_1^T && \text{(2f)}\\
& \lambda^T A_2(u) \leq d_2^T && \text{(2g)}\\
& \lambda \in \mathbb{R}^m && \text{(2h)}\\
& d_1^T x_1 + d_2^T(u_2)x_2 = \lambda^T b(u_2) && \text{(2i)}
\end{aligned}
$$

A short proof of Theorem 1 can be established by invoking the fundamental strong duality theorem in linear programming and its converse (e.g. see [4]).

Although Theorem 1 allows the use of one-level optimization algorithms, note that $\mathcal{P}_2 \in [\mathcal{P}_2]$ is, in general, an instance of a mixed-integer nonlinear program for which global solvers are difficult to implement and relatively inaccessible. Mixed-integer linear algorithms, however, have been extensively studied, and optimized solvers with large-scale capabilities are available (e.g. see [5]). We now derive an exact mixed-integer linear program that produce optimal solutions to instances of $[\mathcal{P}_1]$.

Nonlinear terms in $[\mathcal{P}_2]$ consist of products of continuous primal and dual variables from the lower-level problem, and binary upper-level decision variables. Linearization schemes for such product terms have been previously reported in the mixed-integer programming literature (e.g. see [6]). For example, the vectors u, λ, and x are feasible in (2g) if and only if they satisfy the mixed-integer linear system:

$$\sum_i \sum_k a_{ijk}(\lambda_i - \widetilde{\lambda}_{ijk}) \leq d_{2j},\ \forall j \tag{3a}$$

$$-(1-u_k)\underline{\lambda}_{ijk} \leq \widetilde{\lambda}_{ijk} \leq (1-u_k)\overline{\lambda}_{ijk},\ \forall i,j,k \tag{3b}$$

$$-u_k\underline{\lambda}_{ijk} \leq \lambda_i - \widetilde{\lambda}_{ijk} \leq u_k\overline{\lambda}_{ijk},\ \forall i,j,k \tag{3c}$$

where $\underline{\lambda}_{ijk}$ and $\overline{\lambda}_{ijk}$ are lower and upper bounds, respectively, on the dual variable λ_i.

Note that an equivalent system of equations can similarly be derived for the remaining product terms in (2), thereby yielding a class of mixed-integer linear programs, hereinafter denoted by $[\mathcal{P}_3]$.

Theorem 2: Let $\mathcal{F}(\mathcal{P}_1)$ and $\mathcal{F}(\mathcal{P}_2)$ denote the feasible sets of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Let $\mathcal{O}(\mathcal{P}_1)$ and $\mathcal{O}(\mathcal{P}_2)$ denote the optimal sets of $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Then either of the following conditions hold:

1) $\mathcal{F}(\mathcal{P}_1)$ and $\mathcal{F}(\mathcal{P}_2)$ are empty;
2) $\mathcal{F}(\mathcal{P}_1)$ is nonempty bounded, $\mathcal{F}(\mathcal{P}_2)$ is unbounded, and $\mathcal{O}(\mathcal{P}_3)$ is nonempty bounded.

For some element (x, y), let $\pi_x(x, y)$ denote the projection of (x, y) onto its x-component; that is, $\pi_x(x, y) = x$, and define the projection $\Pi_{S_1}(\mathcal{S}_2)$ of a set $\mathcal{S}_1$ onto a set $\mathcal{S}_2$ by $\Pi_{\mathcal{S}_1}(\mathcal{S}_2) = \{\pi_x(x, y) : x \in S_1, (x, y) \in \mathcal{S}_2\}$.

Theorem 3: A decision vector $(u_0, x_0, \lambda_0) \in \mathcal{U} \times \mathbb{R}^{n+m}$ is feasible in $\mathcal{P}_1$ only if $\pi_x(u_0, x_0, \lambda_0)$ is an extreme point of the polytope $\{x \in \mathbb{R}^n : A_1x_1 + A_2(\pi_{u_2}(u_0))x_2 = b(\pi_{u_2}(u_0)), x \geq 0\}$.

Proposition 1: There exists some real vector $\bar{\lambda} \in \mathbb{R}^m$ with $\|\bar{\lambda}\| < \infty$ such that $\pi_{\mathcal{F}(\mathcal{P}_1)}(\mathcal{O}(\mathcal{P}_3)) = \mathcal{O}(\mathcal{P}_1)$.

Proof: For fixed u, let $x_B(u)$ be an optimal basis to (1d)–(1f) in $\mathcal{P}_1$. From linear programming theory, there exists a permutation matrix $P \in \mathbb{R}^{m\times m}$ such that

$$x(u) = P(u)\widetilde{x}(u), \tag{4a}$$

$$\begin{bmatrix} B(u) & N(u) \end{bmatrix} = A(u)P(u) \tag{4b}$$

$$\tilde{x}(u) = (\tilde{x}_B(u), \tilde{x}_N(u)), \tag{4c}$$

$$\lambda^T(u) = c_B^T(u)B^{-1}(u). \tag{4d}$$

where $B(u)$ is an optimal basis matrix. Note that the existence of a (regular) basis matrix $B(u)$ is guaranteed by Assumption 4. Taking the norm of (4d), we have

$$\|\lambda^T(u)\| = \|c_B^T(u)B^{-1}(u)\| \tag{5a}$$

$$\leq \|c^T(u)\|\|B^{-1}(u)\| \tag{5b}$$

■

Theorem 4 (Projection of $\mathcal{P}_3$ onto $\mathcal{P}_2$): Assume $\overline{\gamma}_{ijk}$, for all i, j, and k, are selected sufficiently large such that for every optimal solution to $\mathcal{P}_3$, $u_k = 1$ only if $-u_k\underline{\lambda}_{ijk} \leq \lambda_i$, and $\lambda_i \leq u_k\overline{\lambda}_{ijk}$, for all i, j, k, are nonbinding. Then $(u^*, x^*, \lambda^*, \widetilde{\lambda}^*)$ is optimal in $\mathcal{P}_3$ entails $(u^*, x^*, \lambda^*, \widetilde{\lambda}^*)$ is optimal in $\mathcal{P}_2$.

It is possible to substitute a product of 0-1 decision variables by an equivalent set of mixed-integer linear equalities or inequalities without any additional integer variables, which leads to the following results.

Theorem 5 (Extension to weak $[\mathcal{P}_1]$): Let weak $[\mathcal{P}_1]$ be the class of Stackelberg games obtained from $[\mathcal{P}_1]$ by weakening the conditions on h, b, c_2, d_2 and A_2 such that each component of the said functions is a sum of monomials in the integer decision variables. Thus weak $[\mathcal{P}_1]$ contains $[\mathcal{P}_1]$. Then, the aforementioned mixed-integer linear transformation is applicable to weak $[\mathcal{P}_1]$.

For practical purposes, it is desirable to obtain tighter bounds on λ_i, for all i. This could be achieved by algebraic schemes for the computation of exact bounds for the solution set of linear interval systems (e.g. [7], [8]), or some modification to an iteration of the simplex algorithm.

III. Conclusion

An approach has been proposed for globally solving a class of nonlinear two-level decision problems, using available mixed-integer linear programming algorithms. Research is ongoing to derive tighter bounds on the linear relaxations of the mixed-integer linear reformulation.

References

[1] S. Dempe, *Foundations of Bilevel Programming*. Boston, MA: Kluwer, 2002.

[2] J. Salmerón, K. Wood, and R. Baldick, "Analysis of electric grid security under terrorist threat," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 905–912, May 2004.

[3] J. M. Arroyo and F. D. Galiana, "On the solution of the bilevel programming formulation of the terrorist threat problem," *IEEE Transactions on Power Systems*, 2005, in press.

[4] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.

[5] R. E. Bixby, M. Fenelon, Z. Gu, E. Rothberg, and R. Wunderling, "MIP: Theory and practice closing the gap," in *System Modelling and Optimization: Methods, Theory and Applications*, M. J. D. Powell and S. Scholtes, Eds. The Netherlands: Kluwer, 2000, pp. 19–50.

[6] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. New York, NY: Oxford Univ. Press, 1995.

[7] C. Janson, "Calculation of exact bounds for the solution of linear interval systems," *Linear Algebra and Its Applications*, vol. 251, pp. 321–340, 1997.

[8] J. Rohn, "Solvability of systems of linear interval equations," *SIAM J. Matrix Anal. Appl.*, vol. 25, no. 1, pp. 237–245, 2003.