Big Data Based Intelligent Decision Support System for Sustainable Regional Development

Hong Zhou
School of Mechanical and Electrical Engineering
Faculty of Health, Engineering and Science
The University of Southern Queensland
Toowoomba, Australia
hong.zhou@usq.edu.au

Christopher Noble, Julie Cotter
Australian Centre for Sustainable Business and Development (ACSBD)
The University of Southern Queensland
christopher.noble, julie.cotter@usq.edu.au

***Abstract*— Timely intelligent decision making is increasingly important for modern society. With the availability of big data and advanced artificial intelligence in decision making, more objective and evidence- based quantitative smart decisions can be made in a timely manner. This research proposed a big data based intelligent decision support system (B-IDSS) for sustainable business development. The system can be used by both the government agencies and corporate business (e.g. farms. mining) in advanced planning, collaboration and management. This paper also addresses the performance optimization as bilevel decision-making problem with one leader and multiple followers. An extended Kuhn-Tucker approach is introduced as one of the algorithms that can be adapted in the system.**

***Keywords*— decision making; business model; sustainability; big data; bilevel programming.**

I. Introduction

Smart and intelligent decision making is crucial for everyone, including individuals, government agencies, non-profit organizations and businesses. Decisions are made on a daily basis or even in a single moment. While some decisions are important others are not, and decisions depend on the actions of others rather than being made independently. Decisions are also subject to the information available and the treatment of that information. In general, the more relevant information acquired, the more knowledge a person can construct, and the smarter the decisions that can be made. With limited information and experience, it is not uncommon that regrettable decisions are made and a big data based intelligent decision making system (B-IDSS) with a reliable algorithm can reduce such costs.

Traditionally decision making is carried out by those duty-bound, e.g. board members of a company or organization. Face-to-face meetings are often held on a regular basis bringing their expertise and knowledge to critical decisions. This process works well for most organizations. However, there are limitations in this traditional decision making process with the increasing frequency that decisions are being made. This face-to-face process is normally very slow and inefficient. The decisions are often made without knowing the consequences to other parties or their reactions (businesses or government). The cost of bring this information to a decision making process and ensuring consistency in the decision process can be high and subject to interference from unforeseen sources.

Many decision making problems involve several interacting organizations or individuals. Often these groups are arranged within a hierarchical structure with independent or perhaps conflicting objectives. The actions of one participant/player affect the choice and payoffs available to the other but neither player has a dominate strategy in the traditional sense. This creates a demand for information to be constantly updated and a reactive decision as another player makes a decision or if the wider decision environment changes.

Game theory has been applied as a useful tool for modelling communications and decision making systems [3]. Modelling of decentralised systems of organizations is undertaken by bilevel programming (BLP) techniques motivated by the game theory of Von Stackelberg [1] from the context of asymmetric economic markets [2]. The complex decision making in hierarchical environment can be regarded as a general non-cooperative bilevel multi-follower decision problem: (1) the upper-level is termed as the leader and the lower-level is termed as the follower. (2) The leader may be able to influence the behaviour of the followers but not completely control the follower's actions. Conversely, the leader may be affected by the follower's behaviour. (3) In this decision system, each decision maker (the leader and followers) tries to optimize his/her own objective by control of his/her own variables. In this way, decisions are made sequentially over time as each player reacts to the decision of another.

Approximately twenty algorithms have been surveyed as part of this research, such as, the K^{th} best approach [4, 5], Kuhn-Tucker approach [6-8], complementarity pivot approach [9], penalty function approach [10, 11], for solving linear BLP problems. The Kuhn-Tucker approach is a valuable analysis tool with a wide range of successful applications for linear BLP [3, 7, 8]. This paper provides an extended Kuhn-Tucker approach as an example to be used in the algorithm component in the proposed system. This algorithm is based on previous work [12-14] on bilevel models by building the capability to analyse an increased range of bilevel decision problems. The other algorithms in optimal decision making can also be included in the proposed B-IDSS framework.

This research proposes a big data based decision support system to assist decision-makers. It integrates state-of-art multi-disciplinary technologies such as data mining, artificial intelligence in decision making, and communications together. The system utilises a significant amount of data in a rich range of forms such as text and multimedia from a wide diversity of sources such as government, non-profit organization and business. The Darling Downs region of Queensland is used as an example where this model may be implemented to cope with issues of local government policy on business sustainability and industry competitiveness. The extended Kuhn-Tucker approach to bilevel programming is chosen as one of the potential algorithms to assist people in different roles make decisions based on another's actions.

The rest of this paper is organized as follows. In Section 2, architecture for B-IDSS system is introduced. Then the sustainability of business development and operation are addressed. An extended Kuhn-Tuck approach which can be adapted in the framework is presented in Section 4. Finally Section 5 concludes the paper with some future research directions.

II. Architecture Of Big Data Based Intelligent Decision Support System (B-IDSS)

This section introduces the proposed architecture for B-IDSS system. The architecture integrates big data resources with a web-based decision support system. Additional components on data/text/multimedia mining are added in the proposed architecture.

The simplified frameworks we developed is shown in Fig. 1 and Fig. 2. Fig. 1 provides an overview of the system with multiple users to access the service through the Internet.

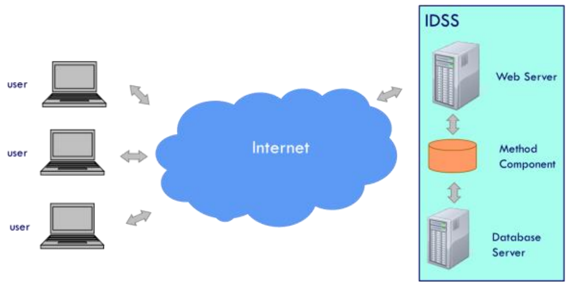

Fig. 1. Architecture of B-IDSS System

Fig. 2 illustrates the major components and interactions among them in the system. It consists of three layers: the top layer is the API (Application Interface) which provides the platform for users to describe the problem, objectives, requirements and other specifications.

The bottom layer is data storage layer, which contains the range of raw big data available, the condensed processed information after mining process, and a database of all the intermediate and final decisions and outcomes as the results of the decisions made by the system.

The layer in the middle is the engine layer. It contains the enabling technologies including data mining, modelling, and the optimization algorithms. Data mining includes text mining, web mining, and multimedia (video/audio) mining (Figure 2). The data mining processes the original data to produce a more concise and useful set of data.

This modelling simulates human values in a mathematical form by defining the utility function based on key economic, environmental and business factors. From the information collected and requirements by the user, a model is chosen to best represent the real situation. Given the model, the most suitable algorithm is chosen to find the best solution. At the end, the solutions are provided and stored in a database for reference of future decision making processes.

The applications of modelling and algorithms have been studied extensively with examples of work in bilevel programming surveyed including [4,5,8,9]. However, with the challenges and opportunities presented by big data, as an application for these models and algorithms, the research area remain active and requires further investigation.

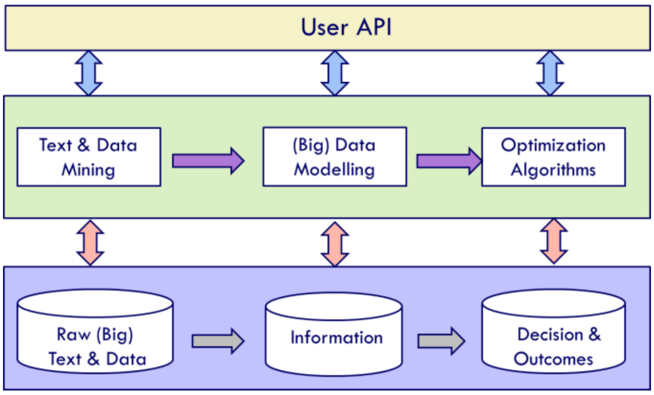

Fig. 2. Framework of B-IDSS System

III. Sustainability of Business Development

With environmental and social change combining with the increasingly competiveness of the global economy, sustainability of businesses [17,18] is a key issue identified for the Darling Downs region [16]. The Darling Downs region in Queensland, Australia, is a prominent food production and mineral resource producing area [19]. Individual business and government organisations need to consider decisions and policies made in light of the interdependence between themselves and the wider community. This complexity calls for new methods or models in the process of decision making in the future to be sustainable. The concept of regional sustainability is illustrated in Fig.3.

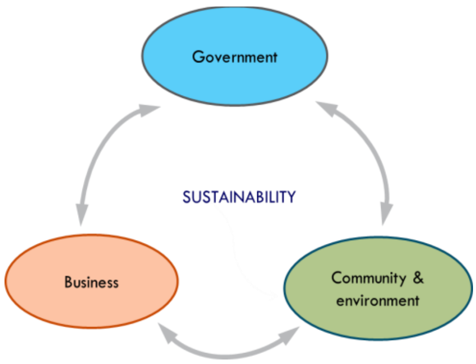

Fig. 3. Regional Sustainability

In this research, the government agency and corporate business are represented by multiple bi-level hierarchical structure as shown in Fig. 4. To maintain the sustainability, each player has two roles, the leader and the follower. The government agency can be regarded as the leader to the businesses (farms and mines) and the follower to the community and environment in the game. On one hand, the government is responsible for industry regulation and planning with the authority to access big data resources from local, interstate and global information networks. On the other hand, the government is also responsible to protect environment and must be supported by the community. An individual business (e.g. farms, corporate farms, mines) is a follower to the government. Meanwhile it acts as a leader to the community and environment in the model. Coalition activity of farms and mines to lobby government is ruled out of this game [20, 21]. Similarly the environment and community as a whole is the follower of the businesses in the region and the leader to the government. The actions of one affect the choice and payoffs available to the other but neither player can completely dominate the other in the traditional sense and sustains itself dynamically.

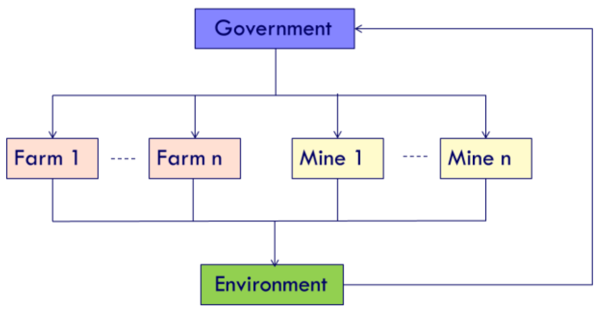

Fig. 4. Hierarchical structure of a sustainable region

IV. The General BLMFP Model and Algorithm

A Bilivel Multi-Follower Programming model (BLMFP) can be applied to address the individual hierarchical links between the government, businesses and environment in the sustainability model shown in Fig.4. In this section, we limit our discussion on the link between the government and businesses. The model and algorithms can be extended to address the other two hierarchical links. The Kuhn-Tucker approach based on the theory and technology of linear bilevel programming is adapted here to provide an example in the proposed B-IDSS. The detailed theory and methods can be found in previous research [12-15].

A. A General BLMFP Model

Let us consider a bilevel game in which there are one government (i.e. regional city council or state government) as a leader and a number of k farms and mines as followers. The government has one set of parameters pertaining to the facility and service it provides (e.g. tax, employment), $x \in X \subset R^n$. Meanwhile the private companies have another set of key parameters which determine the level of profitability and economic performance (e.g. production, price) it offers, that is $y_i \in Y_i \subset R^{m_i}$.

The utility function of the government is defined by $F_i(x, y_i)$ and the utility function of a follower is given by $f_i(x, y_i)$. Therefore we present the model as follows.

For $x \in X \subset R^n, y_i \in Y_i \subset R^{m_i}$ $F : X \times Y_1 \times \ldots \times Y_K \to R^1$, and $f_i : X \times Y_1 \times \ldots \times Y_K \to R^1$, $i = 1,2,\ldots,K$, a linear BLMFP problem, where $K(\geq 1)$ followers are involved, is defined as follows:

$$\min_{x \in X} F(x, y_1, \ldots, y_K) = cx + \sum_{s=1}^{K} d_s y_s \tag{1a}$$

$$\text{subject to } Ax + \sum_{s=1}^{K} B_s y_s \leq b \tag{1b}$$

$$\min_{y_i \in Y_i} f_i(x, y_1, \ldots, y_K) = c_i x + \sum_{s=1}^{K} e_{is} y_s \tag{1c}$$

$$\text{subject to } A_i x + \sum_{s=1}^{K} C_{is} y_s \leq b_i , \tag{1d}$$

where $c \in R^n$, $c_i \in R^n$, $d_i \in R^{m_i}$, $e_{is} \in R^{m_s}$, $b \in R^p$, $b_i \in R^{q_i}$, $A \in R^{p \times n}$, $B_i \in R^{p \times m_i}$, $A_i \in R^{q_i \times n}$, $C_{is} \in R^{q_i \times m_s}$, $i, s = 1,2,\ldots,K$.

Definition 1: A topological space is compact if every open cover of the entire space has a finite subcover. For example, $[a,b]$ is compact in R (the Heine-Borel theorem) [15]. Corresponding to (1), we give following basic definition for linear BLP solution.

Definition 2:

(a) Constraint region of the linear BLP problem:

$$S = \{(x, y_1, \ldots, y_K) \in X \times Y_1 \times \ldots \times Y_k, Ax + \sum_{s=1}^{K} B_s y_s \leq b,$$
$$A_i x + \sum_{s=1}^{K} C_{is} y_s \leq b_i, i = 1,2,\ldots,K\} .$$

The linear BLP problem constraint region refers to all possible combinations of choices that the leader and follower(s) may make.

(b) Projection of S onto the leader's decision space:

$$S(X) = \{x \in X : \exists y_i \in Y_i, Ax + \sum_{s=1}^{K} B_s y_s \leq b,$$
$$A_i x + \sum_{s=1}^{K} C_{is} y_s \leq b_i, i = 1,2,\ldots,K\} .$$

Unlike the rules in non-cooperative game theory where each player must choose a strategy simultaneously, the definition of BLP model requires that the leader moves first by selecting an x in attempt to minimize his objective subjecting to constraints of both upper and lower level.

(c) Feasible set for each follower $\forall x \in S(X)$:

$$S_i(x) = \{y_i \in Y_i : (x, y_1, \ldots, y_K) \in S\}.$$

The feasible region for each follower is affected by the leader's choice of x, and the allowable choices of each follower are the elements of S.

(d) Each follower's rational reaction set for $x \in S(X)$:

$$P_i(x) = \{y_i \in Y_i : y_i \in \arg\min[f_i(x, \hat{y}_i, y_j, j = 1,2,\ldots,K, j \neq i) : \hat{y}_i \in S_i(x)]\}$$

where $i = 1,2,\ldots,K$,

$$\arg\min[f_i(x, \hat{y}_i, y_j, j = 1,2,\ldots,K, j \neq i) : \hat{y}_i \in S_i(x)] = \{y_i \in S_i(x) : f_i(x, y_1, \ldots, y_K) \leq f_i(x, \hat{y}_i, y_j, j = 1,2,\ldots,K, j \neq i), \hat{y}_i \in S_i(x)\}.$$

The followers observe the leader's action and simultaneously react by selecting y_i from their feasible set to minimize their objective functions.

(e) Inducible region:

$$IR = \{(x, y_1, \ldots, y_K) : (x, y_1, \ldots, y_K) \in S, y_i \in P_i(x), i = 1,2,\ldots,K\}.$$

Thus in terms of the above notations, the linear BLP problem can be written as

$$\min\{F(x, y_1, \ldots, y_K) : (x, y_1, \ldots, y_K) \in IR\} \qquad (2)$$

We propose the following theorem to characterize the condition under which there is an optimal solution for a linear BLP problem.

Theorem 1: If S is nonempty and compact, there exists an optimal solution for a linear BLP problem. (Please see [15] for the proof of this theorem.)

B. A General Kuhn-Tucker Approach

Let write a linear programming (LP) as follows:

$$\min f(x) = cx$$
$$\text{subject to } Ax \geq b \qquad x \geq 0$$

where c is an n-dimensional row vector, b an m-dimensional column vector, A an $m \times n$ matrix with $m \leq n$, and $x \in R^n$. Let $\lambda \in R^m$ and $\mu \in R^n$ be the dual variables associated with constraints $Ax \geq b$ and $x \geq 0$, respectively. Bard [2] gave the following proposition.

Proposition 3.1: A necessary and sufficient condition that (x^*) solves above LP is that there exist (row) vectors λ^*, μ^* such that (x^*, λ^*, μ^*) solves:

$$\lambda A - \mu = -c$$
$$Ax - b \geq 0$$
$$\lambda(Ax - b) = 0$$
$$\mu x = 0$$
$$x \geq 0, \lambda \geq 0, \mu \geq 0.$$

Proof:

Let $u_i \in R^p$, $v_i \in R^{q_1+q_2+\ldots+q_K}$ and $w_i \in R^{m_i}$ ($i = 1,2,\ldots,K$) be the dual variables associated with constraints $Ax + \sum_{s=1}^{K} B_s y_s \leq b$, $A'x + \sum_{s=1}^{K} C_s' y_s \leq b'$ and $y_i \geq 0$ $(i = 1,\ldots,K)$, respectively, where $A' = (A_1, A_2, \ldots, A_K)^T$, $C_i' = (C_{i1}, C_{i2}, \ldots, C_{iK})^T$, $b' = (b_1, b_2, \ldots, b_K)^T$. We have a following theorem. (Please see [2] for the details.)

Theorem 2: A necessary and sufficient condition that $(x^*, y_1^*, \ldots, y_K^*)$ solves the linear BLP problem (1) is that there exist (row) vectors $u_1^*, u_2^*, \ldots, u_K^*$, $v_1^*, v_2^*, \ldots, v_K^*$ and $w_1^*, w_2^*, \ldots, w_K^*$ such that $(x^*, y_1^*, \ldots, y_K^*, u_1^*, \ldots, u_K^*, v_1^*, \ldots, v_K^*, w_1^*, \ldots, w_K^*)$ solves:

$$\min_{x \in X} F(x, y_1, \ldots, y_K) = cx + \sum_{s=1}^{K} d_s y_s \qquad (3a)$$

$$\text{subject to } Ax + \sum_{s=1}^{K} B_s y_s \leq b \qquad (3b)$$

$$A'x + \sum_{s=1}^{K} C_s' y_s \leq b' \qquad (3c)$$

$$u_i B_i + v_i C_i' - w_i = -e_{ii} \qquad (3d)$$

$$u_i(b - Ax - \sum_{s=1}^{K} B_s y_s) + v_i(b' - A'x - \sum_{s=1}^{K} C_s' y_s) + w_i y_i = 0 \qquad (3e)$$

$$x \geq 0, y_j \geq 0, u_j \geq 0, v_j \geq 0, w_j \geq 0, \qquad (3f)$$

where $i = 1,2\ldots,K$.

Theorem 2 means that the most direct approach to solving (1) is to solve the equivalent mathematical program given in (3). One advantage this theorem offers is that a more robust model can be solved without introducing new computational difficulties. The detailed proof of this theorem and a numerical example and its results to demonstrate the application of Kuhn-Tucker approach can be found in [15].

V. Conclusion and Future Work

This paper proposed an architecture for big data based intelligent decision support system. The key components are identified and one of the optimization algorithms which can be deployed in the system is introduced. This system provides a software tool which allows government, organizations and businesses to make fast and intelligent decisions. Quantitative modelling and analysis to optimize business performance are possible. The system maximizes the advantage of artificial intelligence through utilising big data resources and helps build sustainable and resilient regions.

Future research includes the construction of a prototype system and case studies by engaging with local farms, mining companies and the regional council. Quantitative modelling and analysis for the optimization of business performance are also among potential research directions.

References

[1] H. Von Stackelberg, The Theory of the Market Economy. Oxford University Press, New York, Oxford, 1952.

[2] J. Bard, Practical Bilevel Optimization: Algorithms and Applications. Kluwer Academic Publishers, USA 1998.

[3] E. Altmana, T. Boulognea, R. El-Azouzia, T. Jiménezb, L.Wynterc, A survey on networking games in telecommunications, Computers & Operations Research vol.33 pp. 286–311, 2006.

[4] W. Candler and R. Townsley, A linear two-level programming problem. Com- puters and Operations Research, vol. 9, pp. 59-76, 1982.

[5] W. Bialas and M. Karwan, Two-level linear programming. Management Science, vol. 30, pp. 1004-1020, 1984.

[6] J. Bard and J. Falk, An explicit solution to the multi-level programming problem. Computers and Operations Research, vol. 9, pp. 77-100, 1982.

[7] W. Bialas and M. Karwan, Multilevel linear programming, Technical Report 78-1, State University of New York at Buffalo, Operations Research Program 1978.

[8] P. Hansen, B. Jaumard, and G. Savard, A new branch-and-bound rules for linear bilevel programming. SIAM Journal on Scientific and Statistical Computing, vol. 13, pp. 1194-1217, 1992.

[9] W. Bialas, M. Karwan, and J. Shaw, A parametric complementary pivot approach for two-level linear programming. Technical Report 80-2, State University of New York at Buffalo, Operations Research Program, 1980.

[10] E. Aiyoshi and K. Shimizu, Hierarchical decentralized systems and its new solution by a barrier method. IEEE Transactions on Systems, Man, and Cybernetics, vol. 11, pp. 444-449, 1981.

[11] D. White and G. Anandalingam, A penalty function approach for solving bi-level linear programs. Journal of Global Optimization, vol. 3, pp. 397-419, 1993.

[12] C. Shi, G. Zhang and J. Lu, On the definition of linear bilevel programming solution. Applied Mathematics and Computation, vol.160, pp. 169-176, 2005.

[13] C. Shi, J. Lu, G. Zhang and H. Zhou, An extended branch and bound algorithm for linear bilevel programming, Applied Mathematics and Computation vol. 180, pp. 529–537, 2006.

[14] J. Lu, C. Shi, G. Zhang and T. Dillon, Model and extended Kuhn-Tucker approach for bilevel multi-follower decision making in a referential-uncooperative situation, International Journal of Global Optimization, vol. 38, pp. 597-608, 2007.

[15] C. Shi, Linear Bilevel Programming Technology – Models and Algorithms, PhD thesis, University of Technology, Sydney, 2005.

[16] Department of State Development, Infrastructure and Planning, Darling Downs regional plan. Brisbane, 2013

[17] T. Lucas, R. Cunningham and G. Lamberton, Small business engagement with sustainability in regional Australia, Journal of Economic and Social Policy, vol. 13, no. 1, art. 3, 2009.

[18] C. Symons and G. Lamberton, Building a social case for business sustainability, Journal of Economic and Social Policy, vol. 16, no. 2, art. 9, 2014.

[19] Department of State Development, Infrastructure and Planning, Darling Downs economic and infrastructure framework. Brisbane, 2013.

[20] D. Schmeidler, The nucleous of a characteristic function game, SIAM Journal on Applied Mathematics, vol. 17, pp. 1163-1170, 1969.

[21] D. Baron and A. Hirsch, Common agency lobbying over coalitions and policy, Economic Theory, vol. 49, pp. 639-681, 2012.