Computers & Industrial Engineering 99 (2016) 260–268

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Optimal network flow: A predictive analytics perspective on the fixed-charge network flow problem

Charles D. Nicholson, Weili Zhang *

School of Industrial and Systems Engineering, University of Oklahoma, USA

ARTICLE INFO

Article history:
Received 2 October 2015
Received in revised form 4 July 2016
Accepted 29 July 2016
Available online 1 August 2016

Keywords:
Network analysis
Fixed charge network flow
Predictive modeling
Critical components

ABSTRACT

The fixed charge network flow (FCNF) problem is a classical NP-hard combinatorial problem with wide spread applications. To the best of our knowledge, this is the first paper that employs a statistical learning technique to analyze and quantify the effect of various network characteristics relating to the optimal solution of the FCNF problem. In particular, we create a probabilistic classifier based on 18 network related variables to produce a quantitative measure that an arc in the network will have a non-zero flow in an optimal solution. The predictive model achieves 85% cross-validated accuracy. An application employing the predictive model is presented from the perspective of identifying critical network components based on the likelihood of an arc being used in an optimal solution.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The fixed charge network flow problem (FCNF) can be easily described as follows. For a given network, each node may have a supply or demand commodity requirement and each incident arc have variable and/or fixed costs associated with commodity flow. The aim of the FCNF is to select the arcs and assign feasible flow to them in order to transfer commodities from supply nodes to demand nodes at a minimal total cost. The transportation problem (Balinski, 1961; El-Sherbiny & Alhamali, 2013), lot sizing problem (Steinberg & Napier, 1980), facility location problem (Aikens, 1985; Daskin, 1995), network design problem (Costa, 2005; Ghamlouche, Crainic, & Gendreau, 2003; Lederer & Nambimadom, 1998) and others (Armacost, Barnhart, & Ware, 2002; Jarvis, Rardin, Unger, Moore, & Schimpeler, 1978) can be modeled as a FCNF.

The FCNF problem is known to be NP-hard (Guisewite & Pardalos, 1990). A significant amount of effort has been invested to study and develop efficient approaches to the FCNF. Many techniques commonly utilize branch and bound to search for an exact solution to the FCNF (Barr, Glover, & Klingman, 1981; Cabot & Erenguc, 1984; Driebeek, 1966; Hewitt, Nemhauser, & Savelsbergh, 2010; Kennington & Unger, 1976; Ortega & Wolsey, 2003; Palekar, Karwan, & Zionts, 1990). Branch and bound however may be inefficient due to lacking tight bounds during the linear relaxation step. Heuristic approaches to find the near-optimal solution of the FCNF have generated considerable research interest (Adlakha & Kowalski, 2010; Antony Arokia Durai Raj, 2012; Balinski, 1961; Kim & Pardalos, 1999; Molla-Alizadeh-Zavardehi, Hajiaghaei-Keshteli, & Tavakkoli-Moghaddam, 2011; Monteiro, Fontes, & Fontes, 2011; Sun, Aronson, McKeown, & Drinka, 1998). State-of-the-art MIP solvers combine a variety of cutting plane techniques, heuristics and the branch and bound algorithm to find the global optimal solution. Modern MIP solvers use preprocessing methods to reduce the search space by taking information from the original formulations, which significantly accelerate the solving processes (Bixby, Fenelon, Gu, Rothberg, & Wunderling, 2000). In this paper, we take a decidedly different approach to leveraging information from the problem formulation and FCNF instances. That is, we are interested in gaining information about how the various topological and component characteristics relate to the selection of arcs used to transmit the optimal flow. At this time, we have found no literature that approaches a study of the FCNF problem from the perspective of statistical learning.

FCNF formulations are useful in many practical problems. Modern societies are heavily dependent on distributed systems, e.g. communication networks (Cohen, Erez, Ben-Avraham, & Havlin, 2000), electric power transmission networks (Dobson, Carreras, Lynch, & Newman, 2007), and transportation networks (Zheng, Gao, & Zhao, 2007). Designing and maintaining such systems is an important research area in network science. In particular, developing resilient network infrastructures (i.e., resilient with respect to natural disasters or intentional attacks) is of utmost importance and the ability to identify critical components in complex networks

* Corresponding author at: School of Industrial and Systems Engineering, University of Oklahoma, 202 W. Boyd St., Room 442, Norman, OK, USA.

E-mail address: weili.zhang-1@ou.edu (W. Zhang).

http://dx.doi.org/10.1016/j.cie.2016.07.030
0360-8352/© 2016 Elsevier Ltd. All rights reserved.

has reached a level of national urgency (Birchmeier, 2007). The destruction or damage of one or more critical components in a networked system could have significant consequences in terms of overall system performance (Bell, 2000; Smith, Qin, & Venkatanarayana, 2003). The definition of component criticality is often associated with an overall network performance metric. A component whose hypothetical failure most impacts the network performance level is identified as critical. A substantial body of work using a variety of methods has focused on identifying critical components within networks, e.g. topological approach (Bompard, Napoli, & Xue, 2009; Crucitti, Latora, & Marchiori, 2005), simulation (Eusgeld, Kröger, Sansavini, Schläpfer, & Zio, 2009), optimization (Bier, Gratz, Haphuriwat, Magua, & Wierzbicki, 2007; Shen, Smith, & Goli, 2012; Zio, Golea, & Rocco, 2012), service measure (Dheenadayalu, Wolshon, & Wilmot, 2004; Scott, Novak, Aultman-Hall, & Guo, 2006) and graph theory (Demšar, Špatenková, & Virrantaus, 2008). In this study we consider an application of our statistical model with respect to identifying critical components wherein the minimum total commodity routing cost, inclusive of fixed costs, is the overall network performance metric.

To the best of our knowledge no existing work has developed models to help characterize predictive network features of optimal solutions to the FCNF. More broadly, little work has been published so far in the application of statistical learning to traditional optimization or network problems. Rocco and Muselli (2004, 2005) developed a decision tree and a hamming clustering model to predict network connectivity reliability in graphs. Hamming clustering is applicable only if both the predicted value and all predictors are binary (Muselli & Liberati, 2002). The binary predictions relating to connectivity were made based on a single type of predictor – the status of each arc in the graph as either failed or operating. Based on this information they attempted to evaluate the reliability of origin-destination connectedness. Empirically they create one network instance (11 nodes, 21 edges) and randomly sample from the possible state space of edge failures. Among the possible 2^{21} states, 2000 were assigned to a training set and 1000 assigned to a test set. The models were developed on the 2000 training observations and highly accurate predictions were observed on the test set. While the predictive models developed were highly accurate, they are inherently linked to the single network instance considered.

In this study we employ a statistical learning technique to analyze the data associated with optimal FCNF solutions and we develop a relatively generalizable model based on several salient network features to predict which arcs will be used in an optimal solution. By solving thousands of generated FCNF instances we collect over 60,000 observations and develop a logistic regression model based on the dataset. This model allows us to quantify the influence of several important network characteristics. The resulting model has several potential applications. In this study, we demonstrate an application for providing an alternative approach to identifying critical network components. The remainder of this paper is organized as follows. Section 2 introduces the background of the FCNF and the logistic regression model. The process for developing the predictive model is discussed in Section 3. The identification of critical components using the model is presented in Section 4. Section 5 summarizes the results and introduces planned future work.

2. Background

2.1. Fixed charge network flow problem

The fixed charge network flow (FCNF) problem is described on a network $G = (N, A)$, where N and A are the set of nodes and arcs, respectively. Let c_{ij} and f_{ij} denote the variable and fixed cost of arc $(i,j) \in A$, respectively. Each node $i \in N$ has a commodity requirement r_i associated with it (if it is a supply node, $r_i > 0$; if a demand node, $r_i < 0$; if a transshipment node, $r_i = 0$). An arc parameter M_{ij} is used in the problem formulation to ensure that the fixed cost f_{ij} is incurred whenever there is a positive flow on arc $(i,j) \in A$. There are two decision variable types: y_{ij} which denotes the decision variable to use arc $(i,j) \in A$ in a solution and x_{ij} denotes the commodity flow on (i,j). The mathematical formulation is as follows,

$$\min \sum_{(i,j)\in A} (c_{ij}x_{ij} + f_{ij}y_{ij}) \tag{1}$$

$$\text{s.t.} \sum_{(i,j)\in A} x_{ij} - \sum_{(j,i)\in A} x_{ji} = r_i \quad \forall i \in N \tag{2}$$

$$0 \leqslant x_{ij} \leqslant M_{ij}y_{ij} \quad \forall (i,j) \in A \tag{3}$$

$$y_{ij} \in \{0,1\} \quad \forall (i,j) \in A \tag{4}$$

Constraint (2) ensures that the inflow and outflow satisfy the supply/demand at node $i \in N$. The parameter M_{ij} in constraint (3) is either the associated arc flow capacity or an artificial arc capacity (for uncapacitated problems). The constraint ensures that the flow on arc $(i,j) \in A$ can be positive only when the arc $(i,j) \in A$ is open ($y_{ij} = 1$). If arc (i,j) does not have a capacity, M_{ij} should be set to a value which is large enough to not inhibit the optimal flow. All problems in this study are uncapacitated and each M_{ij} is set to the total supply in the network. Constraint (4) defines y_{ij} as binary, which makes the problem a 0–1 mixed integer programming problem.

2.2. Logistic regression

Logistic regression is a widely-used technique for classification modeling and is commonly used in business modeling, data mining applications, biological fields, and others (Camdeviren, Yazici, Akkus, Bugdayci, & Sungur, 2007; Hosmer, Lemeshow, & Sturdivant, 2013; Menard, 2002). While there are many classification modeling techniques (e.g., support vector machines, random forests, boosted trees), logistic regression has an advantage regarding model interpretability. Decision trees which are also easy to interpret have a drawback in that they are often unstable. That is, the rules generated by a decision tree are highly sensitive to the instance of training data (Friedman, Hastie, & Tibshirani, 2001). Given that we are interested in analyzing data to understand the characteristics of optimal FCNF solutions, interpretability and stability are important.

We denote the dependent variable (also called a response variable) as Y and define it as follows,

$$Y = \begin{cases} 1, & \text{arc has positive flow in the FCNF optimal solution} \\ 0, & \text{otherwise.} \end{cases}$$

The logistic regression function produces a probability that the response variable equals 1 given the data values observed for the associated k predictor variables, $p_1, \ldots, p_k$,

$$P(Y = 1 | p_1, \ldots, p_k) = \frac{1}{1 + e^{-\left(\beta_0 + \sum_{i=1}^{k} \beta_i p_i\right)}} \tag{5}$$

where the parameters $\beta_1, \ldots, \beta_k$ are regression coefficients determined using maximum likelihood estimation during the modeling process. For a given set of observed values, the binary response variable is set to 1 if the predicted probability exceeds a cut-off point. The details for setting the cut-off value are discussed in Section 3.4.

The β values in a logistic regression model are interpreted similar to linear regression, in that, they represent partial slopes (albeit, the slope is with respect to the log odds) and represent effects assuming all other predictors are held constant. To interpret the regression coefficients in a logistic regression, it is useful to exponentiate the coefficient to compute the odds ratio for that associated predictor. For instance, if predictor p_1 has a regression coefficient β_1, the odds that an arc is selected, $\frac{P(Y=1)}{P(Y=0)}$, is affected by changes in the value of p_1. For every one unit increase in p_1 (holding all other predictors constant) the odds are multiplied by the factor e^{β_1}.

3. Predictive modeling

3.1. Feature engineering

Feature engineering is a term from machine learning used to denote the process of determining and/or deriving predictor variables used in model. Based on initial testing we derive four types of predictors for the classification model: overall network level characteristics, arc specific attributes, linear relaxation based variables, and lastly, variables related to the nodes incident to an arc. These predictors are developed with a basic guiding principle of being easily understandable and accessible from an original problem instance.

Several of the predictors are scaled to range from 0 to 1. As this can help with the interpretation of variables across diverse FCNF instances. To achieve this scaling, various predictors as described below will be divided by either the total supply or the total number of nodes in a generated FCNF instance. Nevertheless, information regarding the actual sizes of the instances are likely to be important as well. Therefore, some variables are not transformed, namely the number of nodes and number of arcs in the instance.

At the network level, the characteristics considered are the number of nodes (n), number of arcs (m), total commodity supply (S), and density of network, $\rho = \frac{m}{n(n-1)}$. Since the total supply amount and the network size are possibly correlated, we transform S to the average supply per node, $\bar{S} = \frac{S}{n}$. The arc specific attributes include the variable cost, the fixed cost, and the ratio of these two values, $\gamma_{ij} = \frac{f_{ij}}{c_{ij}}$ for every $(i,j) \in A$.

The linear relaxation solution of the FCNF problem produces two important predictors for the proposed model. The first of which is related to the value of the flow on arc (i,j) from the relaxed solution (denoted as l_{ij}). Let $\bar{l}_{ij}$ reflect the scaled version of this predictor defined as the ratio of l_{ij} to total available commodity supply S, since S is an upper bound on feasible flow values,

$$\bar{l}_{ij} = \frac{l_{ij}}{S} \quad \forall (i,j) \in A.$$

Note that $0 \leqslant \bar{l}_{ij} \leqslant 1 \ \forall (i,j) \in A$. The second predictor we derive from the linearized solution is the binary variable l^B_{ij} defined as

$$l^B_{ij} = \begin{cases} 1, & \bar{l}_{ij} > 0 \\ 0, & \text{otherwise}. \end{cases}$$

We construct predictors relating to the endpoints of arc (i,j). These predictors include the node types (supply, demand, transshipment), total requirement values, indegrees, and outdegrees for both the tail and head node. The quantity and types of nodes adjacent to i and j may also have an influence on the response variable associated with (i,j). Let t_i be a nominal variable which denotes the type of node $i \in N$ having three factor levels,

$$t_i = \begin{cases} 1, & r_i > 0 \\ 0, & r_i = 0 \\ -1, & r_i < 0 \end{cases}.$$

Let d_{i*} denote the outdegree of node i. Let d^S_{i*} (d^D_{i*}) denote the number of supply (demand) head nodes adjacent to node i. Let r^S_{i*} (r^D_{i*}) denote the sum of supply (demand) requirements of the head nodes adjacent to node i. Similarly, the indegree of node i is denoted as d_{*i}, the number of supply (demand) tail nodes adjacent to node i is denoted by d^S_{*i} (d^D_{*i}), and sum of commodity supply (demand) of tail nodes adjacent to node i are denoted as r^S_{*i} (r^D_{*i}). Fig. 1 illustrates these notations for a small network. The number under each node is the requirement value associated with that node.

To scale the predictors relating to node degrees to the range $[0,1]$, each is divided by the number of nodes n to produce the final predictors $\bar{d}_i$, $\bar{d}_{i*}$, $\bar{d}^S_{i*}$, $\bar{d}^D_{i*}$, $\bar{d}_{*i}$, $\bar{d}^S_{*i}$, $\bar{d}^D_{*i}$ $(\forall i \in N)$. The predictors related with requirements are divided by the total supply S, to produce the variables $\bar{r}_i$, $\bar{r}^S_{i*}$, $\bar{r}^D_{i*}$, $\bar{r}^S_{*i}$, $\bar{r}^D_{*i}$ $(\forall i \in N)$. In summary, the 33 potential predictors for arc selection of an individual arc (i,j) are summarized in Table 1.

3.2. Data collection for model training

In supervised learning, data must be collected on all of the predictors and the associated response variable in order to both train and validate a model. To do so, we generate and solve over a thousand single-commodity FCNF instances using GUROBI 5.6 on a Windows 7 64 bit machine with an Intel Xeon E5-1620 CPU and 8 GB RAM.

FCNF instances are generated by a series of steps to ensure the network is connected and that problem instances are feasible. For the training data, the total number of nodes is randomly generated uniformly between 5 and 15. Each node is labeled with a unique index, $i = 1, 2, \ldots, n$. For every node $i > 1$, an arc (i,k) is added to the instance where $k < i$ is a randomly selected node index. These $n-1$ arcs ensure the network is connected. To randomize the total number of arcs in each final instance, a random number m_{add} is drawn from a uniform discrete distribution on $[0, (n-1)^2]$. Subsequently, m_{add} arcs are added to the instance by iteratively selecting at random two nodes $i,j \in N, \ i \neq j$ which are not adjacent to each other and adding arc (i,j) to the instance. For each arc (i,j) in the instance, arc (j,i) is added to ensure bi-directionality. Therefore, the total quantity of arcs m is an even integer between $2(n-1)$ and $n(n-1)$.

The percentage of supply and demand nodes are each randomly selected on $U(0.15, 0.45)$. For example, in one problem instance, 32% of the nodes might be defined as supply nodes and 18% might be assigned as demand nodes. The requirements for each supply node is randomly assigned as an integer value on $U(1000, 2000)$. The total supply is randomly distributed as negative integer requirements to the demand nodes. The variable costs and fixed costs for each arc are randomly assigned on $U(0, 10)$ and $U(20{,}000, 60{,}000)$, respectively. The test instances referred to in Section 3.4 will be generated in similar fashion with the only exception being the range for determining the number of nodes in each instance. For the training dataset, we solve smaller problems $(5 \leqslant n \leqslant 15)$ so that optimal solutions to many problems can be found quickly. Descriptive statistics of each feature in our training dataset are reported in Table 2. Note that tail and head nodes share the same statistical information. Furthermore, the bi-directionality of the generated instances entails that some of the endpoint-level predictors described in Section 3.1 are equivalent. That is, for the current study, all predictors related to adjacent nodes to an endpoint being either head nodes or tail nodes, must be equivalent, e.g., $\bar{d}_{*i} = \bar{d}_{i*}$ and $\bar{r}_{*i} = \bar{r}_{i*}$, etc. Prior to training the model, the 10 endpoint-level predictors associated with arc (i,j) having subscripts $(*i)$ or $(*j)$ were removed. This leaves 23 potential predictors for model development.

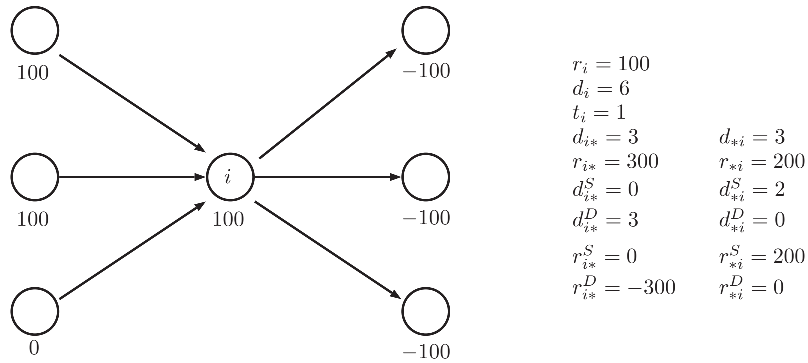

Fig. 1. Predictors associated with node *i*.

Table 1
Candidate predictors for the regression model.

Predictor level	Notation	Description
Network	n	Number of nodes
	m	Total number of arcs
	ρ	Network density
	$\bar{S}$	Average supply
Arc	c_{ij}	Variable cost of arc (i,j)
	f_{ij}	Fixed cost of arc (i,j)
	γ_{ij}	Ratio between fixed cost and variable cost of arc (i,j)
LP solution	$\bar{l}_{ij}$	Normalized value of linearized relaxation solution of arc (i,j)
	l^B_{ij}	Binary variable that equals 1 if $\bar{l}_{ij} > 0$, and 0, otherwise
Endpoints of (i,j)	t_i	Type of tail node i
	t_j	Type of head node j
	$\bar{r}_i$	Requirements of tail node i
	$\bar{r}_j$	Requirements of head node j
	$\bar{r}^S_{i*}$	Sum of supply requirements of head endpoints adjacent to node i
	$\bar{r}^D_{i*}$	Sum of demand requirements of head endpoints adjacent to node i
	$\bar{r}^S_{*i}$	Sum of supply requirements of tail endpoints adjacent to node i
	$\bar{r}^D_{*i}$	Sum of demand requirements of tail endpoints adjacent to node i
	$\bar{r}^S_{j*}$	Sum of supply requirements of head endpoints adjacent to node j
	$\bar{r}^D_{j*}$	Sum of demand requirements of head endpoints adjacent to node j
	$\bar{r}^S_{*j}$	Sum of supply requirements of tail endpoints adjacent to node j
	$\bar{r}^D_{*j}$	Sum of demand requirements of head endpoints adjacent to node j
	$\bar{d}_{i*}$	Outdegree of tail node i
	$\bar{d}^S_{i*}$	Number of supply head nodes adjacent to node i
	$\bar{d}^D_{i*}$	Number of demand head nodes adjacent to node i
	$\bar{d}_{*i}$	Indegree of tail node i
	$\bar{d}^S_{*i}$	Number of supply tail nodes adjacent to node i
	$\bar{d}^D_{*i}$	Number of demand tail nodes adjacent to node i
	$\bar{d}_{j*}$	Outdegree of head node j
	$\bar{d}^S_{j*}$	Number of supply head nodes adjacent to node j
	$\bar{d}^D_{j*}$	Number of demand head nodes adjacent to node j
	$\bar{d}_{*j}$	Indegree of head node j
	$\bar{d}^S_{*j}$	Number of supply tail nodes adjacent to node j
	$\bar{d}^D_{*j}$	Number of demand tail nodes adjacent to node j

Each of the 61,594 observations recorded in the training dataset corresponds to an arc from one of the 1067 FCNF solved instances. The response variable y_{ij} is the binary variable in the optimal solution of the FCNF instances. The number of arcs in the instances ranges from 8 to 210. The instances are each generated such that the ratio of fixed to variable costs are relatively large. The average value for the response variable y_{ij} is quite low: only 11% of arcs across all instances are used in the optimal solutions overall. There is evidence to suggest that statistical learners may be impaired in the presence of highly imbalanced data and there is active research in this area (Chawla, Japkowicz, & Kotcz, 2004; He & Garcia, 2009). Random undersampling on the majority class can be an effective technique for handling imbalance (Burez & Van den Poel, 2009; Hulse, Khoshgoftaar, & Napolitano, 2007) and we apply this method to the training data. The revised dataset for training the model contains 13,349 rows, 51% of which have a value for y_{ij} of 1.

Table 2
Statistical information of features in the training data.

Notation	Min.	1st qu.	Median	Mean	3rd qu.	Max.
n	5.00	10.00	12.00	11.71	14.00	15.00
m	8.00	50.00	84.00	89.86	126.00	210.00
ρ	0.13	0.51	0.72	0.69	0.88	1.00
$\bar{S}$	84.93	307.20	456.75	480.50	622.93	1325.80
c_{ij}	0.00	2.52	5.00	5.01	7.49	10.00
f_{ij}	20,000	29,998	39,983	40,024	50,056	60,000
γ_{ij}	2018	5242	7997	42,839	15,942	141,547,369
$\bar{l}_{ij}$	0.00	0.00	0.00	0.02	0.00	1.00
l^B_{ij}	0.00	0.00	0.00	0.11	0.00	1.00
$t_i,\ t_j$	−1.00	−1.00	0.00	0.15	1.00	1.00
$\bar{r}_i,\ \bar{r}_j$	−1.00	−0.16	0.00	0.00	0.15	1.00
$\bar{r}^S_{i*},\ \bar{r}^S_{j*}$	0.00	0.40	0.60	0.60	0.90	1.00
$\bar{r}^D_{i*},\ \bar{r}^S_{j*}$	−1.00	−1.00	−0.69	−0.65	−0.45	0.00
$\bar{r}^S_{*i},\ \bar{r}^S_{*j}$	0.00	0.00	0.00	0.64	1.01	10.00
$\bar{r}^D_{*i},\ \bar{r}^D_{*j}$	−11.00	−1.00	0.00	−0.66	0.00	0.00
$\bar{d}_{i*},\ \bar{d}_{j*}$	0.07	0.50	0.67	0.64	0.83	0.93
$\bar{d}^S_{i*},\ \bar{d}^S_{j*}$	0.00	0.17	0.27	0.28	0.38	0.83
$\bar{d}^D_{i*},\ \bar{d}^D_{j*}$	0.00	0.10	0.17	0.18	0.25	0.67
$\bar{d}_{*i},\ \bar{d}_{*j}$	0.07	0.50	0.67	0.64	0.83	0.93
$\bar{d}^S_{*i},\ \bar{d}^S_{*j}$	0.00	0.00	0.00	0.28	0.62	0.93
$\bar{d}^D_{*i},\ \bar{d}^D_{*j}$	0.00	0.00	0.00	0.18	0.36	0.93
y_{ij}	0.00	0.00	0.00	0.11	0.00	1.00

3.3. Model development

A logistic regression model is fit based on the balanced training data. We use R version 3.0.2 and the "glm" function from the base statistical package to fit the model using iteratively reweighted least squares for maximum likelihood estimation.

Feature selection is employed by means of backwards stepwise regression based on the Akaike information criterion (AIC) (Akaike, 1974; Venables & Ripley, 2002). Stepwise feature selection techniques are heuristic approaches to find a balance between model complexity and predictive power. *Best subset selection* requires the evaluation of all possible model subsets and is often computationally infeasible. In the present case, best subset selection would require 2^{23} model evaluations.

AIC backward stepwise selection is an iterative procedure that evaluates individual model possibilities by removing predictors one at a time from the current model. Models that are likely to be too complex and "overfit" the training data are penalized. The MASS package for R and the function "stepAIC" are used to perform model selection. Predictors are removed until the AIC score is minimized. The predictors eliminated from the candidate set by the AIC stepwise selection are m, γ_{ij}, $\bar{d}_{i*}$, $\bar{d}_{j*}$, and $\bar{d}^{D}_{j*}$. A predictor may be selected for elimination due to an inherently insignificant contribution to the explanatory power of the statistical model; or possibly, that given the presence of other predictors in the model, the *incremental* predictive power of the eliminated feature is insufficient to justify the additional model complexity from including the feature. The latter is the case in the present model. A separate regression model was fit with only the eliminated variables and each proved to be statistically significant as model covariates. The model estimation and selection converges without issue within 10 min. The final logistic regression model contains 18 predictors. These predictors, their regression coefficients (β), standard errors, odds ratios (OR) computed as e^{β}, and p-values are reported in Table 3. Note that there are two coefficients for the predictors t_i and t_j since these variables each have three factor levels. The reference cases are $t_i = -1$ and $t_j = -1$.

3.4. Model validation

The logistic regression model assigns a probability value to each arc in an FCNF instance. If the probability for arc (i,j) exceeds a threshold value t, the arc is predicted to be used in an optimal solution (i.e., the predicted value for the associated y_{ij} is 1). To maximize observed accuracy, we set the threshold to 0.50. Using 10-fold cross validation the estimated predictive accuracy of the model on the training data set is 0.878. That is, 87.8% of the arcs are correctly predicted as "optimal arcs", those used in an optimal solution, or as "non-optimal arcs", those not used in an optimal solution. The associated confusion matrix associated with the hold-out folds is reported in Table 4. The model correctly predicts 5702 arcs as non-optimal (true negatives are denoted as TN) and 6027 as optimal arcs (true positives denoted as TP). There are 740 false negatives (denoted as FN) and 880 false positives (denoted as FP) in which the model incorrectly assigns a value of 0 to an arc that is an optimal arc and a value of 1 to a non-optimal arc, respectively. Accuracy is defined as the sum of true positives and true negatives divided by the total classifications. The true positive rate (TPR) is defined as $\frac{\mathrm{TP}}{\mathrm{TP+FN}}$ and represents the percentage of optimal arcs captured by the model. The false positive rate (FPR) is defined as $\frac{\mathrm{FP}}{\mathrm{FP+TN}}$ and measures the percentage of non-optimal arcs misclassified as optimal arcs. The threshold value t can be adjusted to reduce false positive rate or false negative rate.

To further validate the model, a set of 182 larger FCNF instances ($15 \leqslant n \leqslant 25, 28 \leqslant m \leqslant 592$) are generated and solved to evaluate the predictive performance for more complex FCNF instances. The test dataset contains 37,651 observations. Similar to the training data, each observation refers to an individual arc. The arcs, their set of associated features, and the observed statuses of use in optimal solutions, are accumulated from all 182 solved FCNF problem instances. Descriptive statistics associated with the test data are reported in Table 5. Since the test data is generated with the same logic as the training data, many of the feature distributions are similar to the training data. The primary difference is with respect to the size of the network instances. To conserve space, only the feature statistics from the test dataset that are notably different from the training data are reported. In the test data, 7% of the arcs are used in optimal solutions. The confusion matrix reported in Table 6 shows that the accuracy remains high at 83.4%. The FPR and TPR values also remain good. This provides evidence that the predictive model is not specific only to the training data, but that it is useful on larger and more complex problem instances.

The receiver operating characteristic (ROC) curve is a common visualization technique for understanding the performance of classification models. The ROC curve plots TPR versus FPR for different threshold values t. The ROC curve associated with the classifier performance on the test data is depicted in Fig. 2. The area under the ROC curve (AUC) ranges from 0 to 1 and is a measure of the discriminatory power of the model, with a perfect predictive model obtaining a value of 1. The diagonal line on the chart represents a baseline model with no discriminatory power and has an AUC of 0.5. AUC values exceeding 0.90 denote models with "outstanding discrimination" performance (Hosmer et al., 2013). The AUC value obtained with the logistic regression predictive model is 0.95.

To explore the model performance on the test data in more detail, the AUC performance metric is reported in Table 7 for subsets of the test data associated with various instance characteristics. In particular, for problem instances with high/low quantities of nodes, arcs, density values, and average supplies per node are isolated and evaluated. The distinction between "high" and "low" is made with respect to the median value of the corresponding feature as reported in Table 5. The number of observations associated with each subset specification is also reported. The two character-

Table 3
Final logistic regression model.

Predictors	β	Std. error	OR	Pr(> |z|)
(Intercept)	9.432	0.288	12481.465	< 0.001
n	−0.131	0.011	0.877	< 0.001
ρ	−2.451	0.255	0.086	< 0.001
$\bar{S}$	−0.0002	< 0.001	1.000	0.080
c_{ij}	−0.087	0.010	0.917	< 0.001
f_{ij}	−0.0001	< 0.001	1.000	< 0.001
$\bar{l}_{ij}$	5.503	0.513	245.427	< 0.001
$\bar{l}^{B}_{ij}$	1.432	0.116	4.187	< 0.001
$t_i = 0$	−2.142	0.130	0.117	< 0.001
$t_i = 1$	1.226	0.132	3.408	< 0.001
$t_j = 0$	−3.423	0.134	0.033	< 0.001
$t_j = 1$	−1.062	0.131	0.346	< 0.001
$\bar{r}_i$	0.818	0.216	2.266	< 0.001
$\bar{r}_j$	−0.464	0.209	0.629	0.0262
$\bar{r}^{S}_{i*}$	0.997	0.172	2.710	< 0.001
$\bar{r}^{D}_{i*}$	0.844	0.157	2.326	< 0.001
$\bar{r}^{S}_{j*}$	−1.035	0.173	0.355	< 0.001
$\bar{r}^{D}_{j*}$	−0.638	0.122	0.529	< 0.001
$\bar{d}^{S}_{i*}$	−3.021	0.416	0.049	< 0.001
$\bar{d}^{D}_{i*}$	−1.088	0.441	0.337	0.014
$\bar{d}^{S}_{j*}$	1.286	0.413	3.618	0.002

Note: i refers to the tail node and j refers to the head node of arc $(i,j) \in A$.

Table 4
Cross-validated confusion matrix with training data.

	Predicted		
	$y_{ij} = 0$	$y_{ij} = 1$	
Actual			
$y_{ij} = 0$	5702 (TN)	880 (FP)	FPR: 0.13
$y_{ij} = 1$	740 (FN)	6027 (TP)	TPR: 0.89

Table 5
Notable feature values in the test dataset.

Feature	Min.	1st qu.	Median	Mean	3rd qu.	Max.
n	15.00	19.00	20.00	20.62	23.00	25.00
m	28.00	192.00	268.00	271.90	356.00	592.00
ρ	0.10	0.51	0.68	0.66	0.87	0.99
$\bar{S}$	58.22	312.20	476.90	464.20	565.50	993.60
y_{ij}	0.00	0.00	0.00	0.07	0.00	1.00

Table 6
Confusion matrix with test data.

	Predicted		
	$y_{ij} = 0$	$y_{ij} = 1$	
Actual			
$y_{ij} = 0$	29,100 (TN)	6083 (FP)	FPR: 0.17
$y_{ij} = 1$	157 (FN)	2310 (TP)	TPR: 0.94

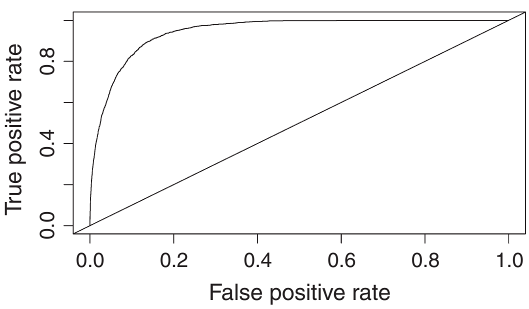

Fig. 2. ROC plot for test dataset.

istics that stand out are the number of arcs and graph density. In both cases, the results indicate that the model has potentially better discriminatory power for instances with a larger number of arcs and increased density.

3.5. Model interpretation

The interpretation of the regression coefficients in Table 3 reveals several interesting relationships between the characteristics of FCNF instances and the arcs used in optimal solutions. Several of these relationships are now discussed.

3.5.1. Network-level predictors

The coefficients for the number of nodes, n, and the graph density, ρ, are both negative. This is a logical outcome. That is, as the size and density of the network increases, the likelihood that a given arc is chosen in an optimal solution decreases. Intuitively, increasing the density results in more alternative paths in the network and therefore, the overall probability for an arc to be selected in the optimal solution is reduced. The average supply per node, $\bar{S}$ also has a small but negative coefficient.

3.5.2. Arc-level predictors

At the arc level, the regression coefficient for variable cost c_{ij} is also negative and has an odds ratio equal to 0.917. This means that for every 1 unit increase in c_{ij} (all other factors held constant) the odds that arc (i,j) is used in an optimal solution decreases by 8.3% where percent change in odds is computed as $100\% \times (e^{\beta} - 1)$.

The coefficient for fixed cost is also negative. While the absolute value of the fixed cost predictor coefficient is much smaller than the variable cost predictor coefficient, it is important to consider the measurement of scale. The range of values used in the instance data for the variable cost and the fixed costs are vastly different, i.e., in the training data, $0 \leqslant c_{ij} \leqslant 10$, whereas $20{,}000 \leqslant f_{ij} \leqslant 60{,}000$. A single unit increase in fixed cost for a given arc does not make a notable impact on the selection probability, however, since the range of values is much larger, a large increase in fixed cost is possible.

While the odds ratios provided in Table 3 correspond to a one unit change increase in the predictor value, the odds ratio for a $k > 0$ increase of the predictor value is easily computed as $e^{k\beta}$ and the corresponding percent change in odds is $100\% \times (e^{k\beta} - 1)$. For example, with all other factors held constant, if an arc with fixed cost originally equal to 40,000 has its cost increased by 10% (to 44,000), then the odds of selection decrease by 44%. For the variable cost, however, a similar 10% increase (from 4 to 4.4) would only decrease the selection odds by 3.4%.

3.5.3. LP solution-level predictors

Notably, the coefficients related to the LP solution predictors are both positive. An arc which has positive flow in the linear relaxed FCNF ($\bar{I}^{B}_{ij} = 1$) has 319% increase in odds for selection as compared to an arc which is not used in the LP relaxed solution. The associated standardized flow predictor, $\bar{l}_{ij}$, has value ranging from 0 to 1. A value of 1 implies that the commodity flow is at maximum capacity. For the problems under study, the arcs are uncapacitated and the logical upper limit to flow on (i,j) is M_{ij} which equals the total supply in the instance. For every ten percentage point increase in this value, arc selection odds increase by 73%.

Traditionally, information related to LP relaxation solution values are commonly employed to guide the branch and bound technique at either the root node or branching nodes (Danna, Rothberg, & Le Pape, 2005; Fischetti & Lodi, 2003; Lodi, 2010; Zhang & Nicholson, submitted for publication). While these two predictors are highly influential in the present model, they are insufficient by themselves to accurately predict arc selection. A logistic regression model built solely from linear relaxation information resulted in diminished model accuracy, with the reduced performance coming from the revised model's inability to identify true positives (TPR reduces from 0.89 to 0.58): the simpler model produces many more false negatives than the full model (2844 compared to 740). Additionally, the model based only on the LP relaxation features produces an AUC value of 0.71 which is on the very low end of "acceptable discrimination" (Hosmer et al., 2013).

3.5.4. Endpoint-level predictors

Finally, characteristics associated with the endpoints of an arc are notably significant factors in accurately predicting that an arc will be selected for use in an optimal solution. An arc is more likely used in an optimal solution if its tail node is a supply node as opposed to a demand node. In particular, the odds for selection of arc (i,j) increase by 241% when $t_i = 1$ (supply) as opposed to

Table 7
Model performance for various test instance characteristics.

	Nodes		Arcs		Density		Average supply	
	$n > 20$	$n \leqslant 20$	$m > 268$	$m \leqslant 268$	$\rho > 0.68$	$\rho \leqslant 0.68$	$\bar{S} > 476.9$	$\bar{S} \leqslant 476.9$
Obs	18,658	18,992	18,770	18,880	18,912	18,738	18,728	18,922
AUC	0.949	0.951	0.968	0.934	0.972	0.931	0.949	0.948

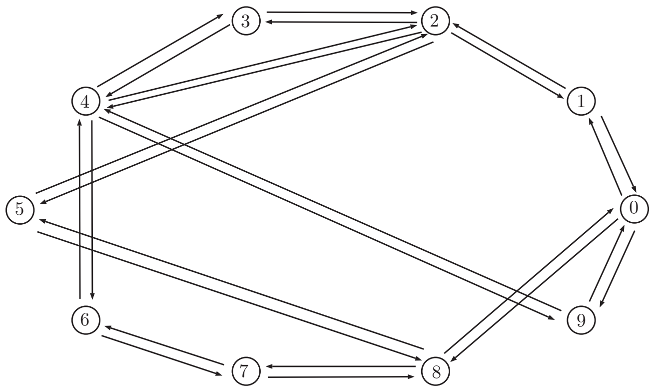

Fig. 3. Small directed network with 10 nodes and 28 arcs.

Table 8
Probability and failure effect for each arc in network.

Arc(i,j)	c_{ij}	f_{ij}	CIM_{ij}	η_{ij}
(7,6)	7.20	41011.33	0.99	10.22 %
(9,0)	6.51	44846.94	0.99	6.78 %
(1,0)	9.69	42158.06	0.99	7.81 %
(9,4)	5.95	22028.81	0.97	7.81 %
(6,7)	1.93	24561.2	0.97	10.22 %
(0,9)	3.10	28765.87	0.95	6.78 %
(1,2)	7.17	44364.01	0.94	0.37 %
(2,5)	8.90	52552.74	0.86	0.37 %
(8,0)	7.54	26588.18	0.68	4.54 %
(5,8)	3.90	28365.83	0.67	7.81 %
(7,3)	1.47	43307.38	0.57	7.81 %
(4,6)	9.09	53894.9	0.53	4.54 %
(8,7)	0.13	33349.63	0.50	7.81 %
(3,7)	4.93	40543.16	0.41	0.37 %
(4,9)	1.21	48247.28	0.26	0.00 %
(7,8)	5.68	49949.96	0.23	0.00 %
(5,2)	2.53	46060.27	0.23	0.00 %
(0,8)	8.61	47039.03	0.17	0.00 %
(8,5)	4.01	57456.67	0.16	0.00 %
(2,1)	5.07	50998.22	0.16	0.00 %
(6,4)	7.67	48589.75	0.12	0.00 %
(2,3)	5.90	31594.21	0.08	0.00 %
(3,2)	3.43	41127.17	0.06	0.00 %
(0,1)	1.27	56093.72	0.04	0.00 %
(4,2)	7.62	39342.56	0.04	0.00 %
(4,3)	4.77	46389.37	0.02	0.00 %
(3,4)	6.18	48528.36	0.02	0.00 %
(2,4)	3.63	54166.87	0.01	0.00 %

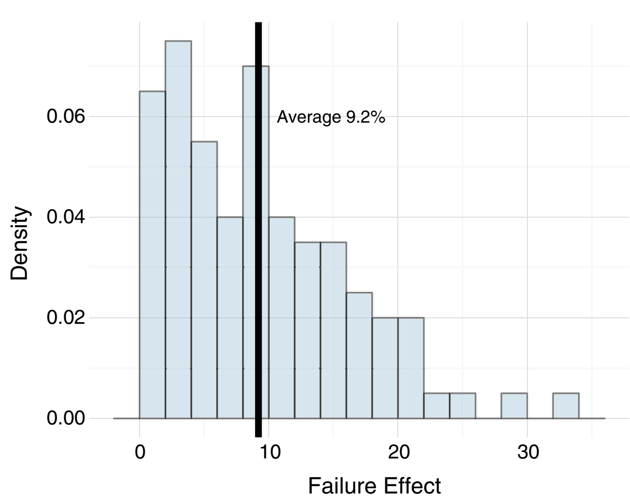

Fig. 4. Distribution of failure effect by removing top two critical arcs.

the reference class, $t_i = -1$ (demand). There is a 65% decrease in odds, however, if the head node is a supply node ($t_j = 1$) instead of a demand node ($t_j = -1$). The coefficients associated with the head node or tail node of (i,j) being transhipment nodes (i.e., $t_i = 0$ and $t_j = 0$) are both highly negative. This implies a notable decrease in likelihood of arc selection with respect to the associated reference class.

Each unit increase in the standardized requirements at a tail node ($\bar{r}_i$) increases the odds of arc selection by a factor of 2.27. Each unit increase of requirements at a head node ($\bar{r}_j$) decreases the odds of selecting arc (i,j) by 37%. Furthermore, as the total supply or total demand at nodes adjacent to the tail node i increases ($\bar{r}_i^S$ or $\bar{r}_i^D$), the odds of (i,j) selection increase by 171% and 133%, respectively. Interestingly, this the opposite effect is observed when considering the head node. As the total supply or total demand at nodes adjacent to the head node j increases ($\bar{r}_j^S$ or $\bar{r}_j^D$), the likelihood of (i,j) selection decreases.

The number of supply or demand nodes adjacent to the tail node i of arc (i,j) also is important. Holding all other factors constant, an increase in either $\bar{d}_{i*}^S$ or $\bar{d}_{j*}^D$ decreases the likelihood of arc selection. This is an interesting finding. Consider, for example the tail node of (i,j), holding $\bar{r}_i^S$ and other factors constant, the model gives a higher likelihood of selection if i is adjacent to a fewer number of nodes with large supply as opposed to a larger quantity of nodes with smaller supplies. For every extra supply node adjacent to i, the odds of selecting (i,j) decreases by 95%. For the head node, an increase in number of adjacent supply nodes increases the odds of selection.

4. Critical components identification

The logistic regression model successfully discriminates between "optimal" and "non-optimal" arcs in FCNF solutions. In this section we develop and demonstrate an application of such information for critical network component identification. A component importance measure (CIM) is often computed to rank nodes or arcs in terms of their potential impact on a network performance measure. The performance measure we use is the FCNF optimal objective value. Since the FCNF problem is NP-hard, admittedly no existing literature uses it as a measure for network performance. However, the optimal objective values implicitly encompasses a variety of metrics which are widely used to evaluate network performance, e.g. network flow, link capacity, delivery costs, and network connectivity. The predictive model provides an efficient and practical method to produce importance measures related to the FCNF optimal solution. We define the CIM as the arc probability values assigned by the classifier. Intuitively, destruction of the more probable "optimal arcs" will likely result in a non-optimal (more expensive) routing of commodity flow.

The novel CIM is implemented and evaluated on a small FCNF instance. We define the actual criticality of an arc as the percentage increase in the FCNF solution cost caused by the destruction of the arc. Let z_0 denote the FCNF solution cost with all arcs present and undamaged. Let z_{ij} denote the optimal objective when arc (i,j) is removed from the network and let η_{ij} denote the associated failure effect,

$$\eta_{ij} = \frac{z_{ij} - z_0}{z_0} \times 100\%.$$

While it is not practical to compute to η_{ij} for real-world FCNF problems, the predictive model is computationally efficient. For the CIM to be valid, there should be a strong rank correlation between CIM and the η values.

To illustrate the method, consider the 10 node FCNF instance in Fig. 3. Nodes 1, 7 and 9 are supply nodes and nodes 0, 5 and 6 are demand nodes. The arrow on the line indicates the direction and each node-pair has two directed arcs, for a total of 28 arcs. The value of z_0 for this instance is 316,074. Table 8 displays the variable cost, fixed cost, CIM, and failure effect for each arc comprising the network. Note, since this is a directed network, arc $(i,j) \in A$ is not the same as arc $(j,i) \in A$ and neither are the failure effects. For example, the CIM value and failure effect of arc $(9,4)$ are much higher than arc $(4,9)$.

Based on this table, the CIM of arcs $(7,6)$ and $(9,0)$ is 0.99, and their failure effects are 10.22% and 6.78%, respectively. Although they do not have the lowest variable cost or fixed cost, the endpoints of these two arcs are a pair of supply and demand nodes. Some arcs with low costs, e.g. $(6,7), (0,9), (8,0), (5,8)$ and $(8,7)$, have high CIM values as well. These arcs can be contrasted with arcs with very low CIM values, which do not have any failure effects. The Spearman's rank correlation between the CIM and the failure effect values is 0.87.

In order to evaluate the quality of the CIM, we conduct an experiment on 100 generated FCNF problems with a variety of densities. Each instance is composed of 20 nodes with the number of arcs ranging from 38 to 380. In this experiment, we perform a comparison of the failure effects resulting from the destruction (i.e., removal) of the two arcs with the highest CIM values as compared to impact from destroying the two arcs with the lowest CIM values.

The failure effect distribution resulting from the removal of the top critical arcs is displayed in Fig. 4. The average effect due to the failures of the most critical arcs is a 9.2% increase in the objective value. In some cases, the failure effect exceeds 30%. The failure effect resulting from the removal of the two lowest CIM rated arcs averages only 0.21% and in 94% of the test cases, the failure effect is 0%. Consequently, it is reasonable to conclude that the arcs with the highest CIM values are more important to the FCNF solutions and may deserve more investment to protect them from various hazards.

5. Conclusions

In this investigation we develop a predictive model to determine whether or not arcs are selected for flow in an optimal solution of a FCNF problem. To do so, we generate and solve over 1000 FCNF instances. The final model, based on 18 derived network related features, allows for high quality discrimination of "optimal" and "non-optimal" arcs. Application to larger FCNF instances retain the predictive performance.

Since we employ a logistic regression technique, the model also has useful explanatory power regarding the predictors defined. The model reveals important quantitative information regarding why arcs are selected in FCNF solutions. Predictors relating to the type, degree, and requirements of the tail and head nodes of an arc are highly relevant to arc selection. The linear relaxation information, while inadequate alone, works well within the larger model to identify optimal arcs. The model predictors and their regression coefficients are rational and intuitive. It is believed that these relationships are valid, at least directionally, among a wider class of FCNF instances and can be used as a "rule-of-thumb" to help understand the problem class. The successful application of the model to the larger problem instances supports this hypothesis.

One successful application of the predictive model is presented in this work. A components importance index is developed based on the probability of arc selection. Empirically, we demonstrate that the impact on network performance from failures of high CIM arcs is substantially greater than that from failures of low CIM arcs.

Since the FCNF is an NP-hard problem with widespread application, the problem is, and has been, the focus of many research efforts for many years. This paper provides an original method to analyze the optimal solution of the FCNF. In future work, we will continue to explore applications of this model in areas such as reducing computational time for exact solutions and developing novel approximate approaches for the FCNF. Furthermore, we realize that there is significant future work in exploring model development to improve predictive accuracy and/or generalization. Additionally, there is much to learn regarding predictive modeling relating to distinct topological patterns and more complex network instances. We hope to continue to develop and motivate ideas in this area.

References

Adlakha, V., & Kowalski, K. (2010). A heuristic algorithm for the fixed charge problem. *Opsearch, 47*(2), 166–175.

Aikens, C. (1985). Facility location models for distribution planning. *European Journal of Operational Research, 22*(3), 263–279.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

Antony Arokia Durai Raj, K., & Rajendran, C. (2012). A genetic algorithm for solving the fixed-charge transportation model: Two-stage problem. *Computers & Operations Research, 39*(9), 2016–2032.

Armacost, C., Barnhart, A., & Ware, K. (2002). Composite variable formulations for express shipment service network design. *Transportation Science, 36*(1), 1–20.

Balinski, M. (1961). Fixed-cost transportation problems. *Naval Research Logistics Quarterly, 8*(1), 41–54.

Barr, R., Glover, F., & Klingman, D. (1981). A new optimization method for large scale fixed charge transportation problems. *Operations Research, 29*(3), 448–463.

Bell, M. (2000). A game theory approach to measuring the performance reliability of transport networks. *Transportation Research Part B: Methodological, 34*(6), 533–545.

Bier, V., Gratz, E., Haphuriwat, N., Magua, W., & Wierzbicki, K. (2007). Methodology for identifying near-optimal interdiction strategies for a power transmission system. *Reliability Engineering & System Safety, 92*(9), 1155–1161.

Birchmeier, J. (2007). Systematic assessment of the degree of criticality of infrastructures. In *Proceedings of ESREL* (pp. 859–864).

Bixby, E., Fenelon, M., Gu, Z., Rothberg, E., & Wunderling, R. (2000). MIP: Theory and practice closing the gap. In *System modelling and optimization* (pp. 19–49). Springer.

Bompard, E., Napoli, R., & Xue, F. (2009). Analysis of structural vulnerabilities in power transmission grids. *International Journal of Critical Infrastructure Protection, 2*(1), 5–12.

Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. *Expert Systems with Applications, 36*, 4626–4636.

Cabot, A., & Erenguc, S. (1984). Some branch-and-bound procedures for fixed-cost transportation problems. *Naval Research Logistics Quarterly, 31*(1), 145–154.

Camdeviren, H., Yazici, A., Akkus, Z., Bugdayci, R., & Sungur, M. (2007). Comparison of logistic regression model and classification tree: An application to postpartum depression data. *Expert Systems with Applications, 32*(4), 987–994.

Chawla, N., Japkowicz, N., & Kotcz, A. (2004). Editorial: Special issue on learning from imbalanced data sets. *ACM Sigkdd Explorations Newsletter, 6*(1), 1–6.

Cohen, R., Erez, K., Ben-Avraham, D., & Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical Review Letters, 85*(21), 4626.

Costa, A. (2005). A survey on benders decomposition applied to fixed-charge network design problems. *Computers & Operations Research, 32*(6), 1429–1450.

Crucitti, P., Latora, V., & Marchiori, M. (2005). Locating critical lines in high-voltage electrical power grids. *Fluctuation and Noise Letters, 5*(02), L201–L208.

Danna, E., Rothberg, E., & Le Pape, C. (2005). Exploring relaxation induced neighborhoods to improve MIP solutions. *Mathematical Programming, 102*(1), 71–90.

Daskin, M. (1995). *Network and discrete location: Models, algorithms, and applications*. John Wiley & Sons.

Demšar, U., Špatenková, O., & Virrantaus, K. (2008). Identifying critical locations in a spatial network with graph theory. *Transactions in GIS, 12*(1), 61–82.

Dheenadayalu, Y., Wolshon, B., & Wilmot, C. (2004). Analysis of link capacity estimation methods for urban planning models. *Journal of Transportation Engineering, 130*(5), 568–575.

Dobson, I., Carreras, B., Lynch, V., & Newman, D. (2007). Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 17*(2), 026103.

Driebeek, N. (1966). An algorithm for the solution of mixed integer programming problems. *Management Science, 12*(7), 576–587.

El-Sherbiny, M., & Alhamali, R. (2013). A hybrid particle swarm algorithm with artificial immune learning for solving the fixed charge transportation problem. *Computers & Industrial Engineering, 64*(2), 610–620.

Eusgeld, I., Kröger, W., Sansavini, G., Schläpfer, M., & Zio, E. (2009). The role of network theory and object-oriented modeling within a framework for the

vulnerability analysis of critical infrastructures. *Reliability Engineering & System Safety, 94*(5), 954–963.
Fischetti, M., & Lodi, A. (2003). Local branching. *Mathematical Programming, 98*(1–3), 23–47.
Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer.
Ghamlouche, I., Crainic, T., & Gendreau, M. (2003). Cycle-based neighbourhoods for fixed-charge capacitated multicommodity network design. *Operations Research, 51*(4), 655–667.
Guisewite, G., & Pardalos, P. (1990). Minimum concave-cost network flow problems: Applications, complexity, and algorithms. *Annals of Operations Research, 25*, 75–99.
He, H., & Garcia, E. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284.
Hewitt, M., Nemhauser, G., & Savelsbergh, M. (2010). Combining exact and heuristic approaches for the capacitated fixed-charge network flow problem. *INFORMS Journal on Computing, 22*(2), 314–325.
Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied logistic regression* (3rd ed.). New York: John Wiley & Sons, Inc..
Hulse, J., Khoshgoftaar, T., & Napolitano, A. (2007). Experimental perspectives on learning from imbalanced data. In *Proceedings of the 24th international conference on machine learning* (pp. 935–942).
Jarvis, J., Rardin, R., Unger, V., Moore, R., & Schimpeler, C. (1978). Optimal design of regional wastewater systems: A fixed-charge network flow model. *Operations Research, 26*(4), 538–550.
Kennington, J., & Unger, E. (1976). A new branch-and-bound algorithm for the fixed-charge transportation problem. *Management Science, 22*(10), 1116–1126.
Kim, D., & Pardalos, P. (1999). A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure. *Operations Research Letters, 24*, 195–203.
Lederer, P., & Nambimadom, R. (1998). Airline network design. *Operations Research, 46*(6), 785–804.
Lodi, A. (2010). Mixed integer programming computation. In *50 years of integer programming 1958-2008* (pp. 619–645). Springer.
Menard, S. (2002). *Applied logistic regression analysis* (Vol. 106). Sage.
Molla-Alizadeh-Zavardehi, S., Hajiaghaei-Keshteli, M., & Tavakkoli-Moghaddam, R. (2011). Solving a capacitated fixed-charge transportation problem by artificial immune and genetic algorithms with a Prüfer number representation. *Expert Systems with Applications, 38*(8), 10462–10474.
Monteiro, M., Fontes, D., & Fontes, F. (2011). An ant colony optimization algorithm to solve the minimum cost network flow problem with concave cost functions. In *GECCO* (pp. 139–146).
Muselli, M., & Liberati, D. (2002). Binary rule generation via hamming clustering. *IEEE Transactions on Knowledge and Data Engineering, 14*(6), 1258–1268.
Ortega, F., & Wolsey, L. (2003). A branch-and-cut algorithm for the single-commodity, uncapacitated, fixed-charge network flow problem. *Networks, 41*(3), 143–158.
Palekar, U., Karwan, M., & Zionts, S. (1990). A branch-and-bound method for the fixed charge transportation problem. *Management Science, 36*(9), 1092–1105.
Rocco, C., & Muselli, M. (2004). Empirical models based on machine learning techniques for determining approximate reliability expressions. *Reliability Engineering & System Safety, 83*(3), 301–309.
Rocco, C., & Muselli, M. (2005). Approximate multi-state reliability expressions using a new machine learning technique. *Reliability Engineering & System Safety, 89*(3), 261–270.
Scott, D., Novak, D., Aultman-Hall, L., & Guo, F. (2006). Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography, 14*(3), 215–227.
Shen, S., Smith, J., & Goli, R. (2012). Exact interdiction models and algorithms for disconnecting networks via node deletions. *Discrete Optimization, 9*(3), 172–188.
Smith, B., Qin, L., & Venkatanarayana, R. (2003). Characterization of freeway capacity reduction resulting from traffic accidents. *Journal of Transportation Engineering, 129*(4), 362–368.
Steinberg, E., & Napier, H. (1980). Optimal multi-level lot sizing for requirements planning systems. *Management Science, 26*(12), 1258–1271.
Sun, M., Aronson, J., McKeown, P., & Drinka, D. (1998). A tabu search heuristic procedure for the fixed charge transportation problem. *European Journal of Operational Research, 106*(2), 441–456.
Venables, W., & Ripley, B. (2002). *Modern applied statistics with S*. Springer.
Zhang, W., & Nicholson, C. (2015). Objective scaling ensemble approach for integer linear programming (submitted for publication).
Zheng, J., Gao, Z., & Zhao, X. (2007). Clustering and congestion effects on cascading failures of scale-free networks. *EPL (Europhysics Letters), 79*(5), 58002.
Zio, E., Golea, L., & Rocco, C. (2012). Identifying groups of critical edges in a realistic electrical network by multi-objective genetic algorithms. *Reliability Engineering & System Safety, 99*, 172–177.