Using Machine Learning to Balance Energy Cost and Emissions in Optical Networks

Arash Deylamsalehi, Dylan A. P. Davis, Pegah Afsharlar, Mehdi Bahrami, Wei-Peng Chen, and Vinod M. Vokkarane

***Abstract*—The goal of energy cost-aware routing and wavelength assignment (RWA) is to minimize the total electricity expenditure in an optical network. While effective, simply aiming to reduce the electricity consumed does not necessarily mitigate the environmental impact. A new approach is required to simultaneously reduce both the electricity cost and emissions produced as a by-product of RWA. We present a method for doing so through the use of mixed-integer linear programming (MILP), which can find the optimal solution that minimizes the electricity cost of RWA for a static set of requests while keeping emissions under a specified cap. This objective is quantitatively compared to alternative goals, including directly minimizing the emissions produced, reducing the length of established paths, and balancing reductions in both emissions and electricity cost simultaneously. As MILP computation is costly and the results are required in near real time to react to changing prices, we present a solution that employs a well-known supervised machine learning algorithm, logistic regression, that predicts the cost saving paths in real time. Using a dataset of 808,024 records (70% for training and 30% for testing) output from the MILP, we find that this logistic regression model predicts the most cost-efficient path with 92.5% accuracy.**

***Index Terms*—Green networking; Real-time pricing; Routing and wavelength assignment; Unicast.**

I. Introduction

Core networks consume about 20 TWh of energy annually. They are responsible for the release of 11 megatons of carbon dioxide emissions and $1.2B of electricity expenditures [1]. With the growth in cloud computing, Internet of Things, and other bandwidth-hungry applications (e.g., 4KTV), the scale of emissions produced and dollars spent has been increasing dramatically. The increase in energy consumption has sparked interest in attempts to improve energy efficiency. Traditionally, it has been assumed that improving efficiency is sufficient for minimizing total overall energy consumption. However, studies have shown that higher efficiency can actually lead to higher consumption [2]. Therefore, to truly reduce consumption, core network operators must focus on efforts to directly decrease the amount of energy spent through changing how demands are routed through the network.

Today's core communication networks span cross-country or cross-continental regions. Consequently, the network operators distribute their infrastructure across different geographical locations worldwide to improve their reliability and quality of service. Such large and often diverse areas might be served by multiple competing power markets throughout, as shown with the NSFnet topology depicted in Fig. 1. The electricity price and fuel mix differ from site to site at any given time and over time at any given site. Over time, the cost can increase or decrease by up to a factor of 11 [3]. Thus, shifting a partial amount of electricity consumption from a given location to other alternative locations leads to a significant potential for reducing the electricity cost and/or emissions of any consumer.

In recent years, power markets have started providing information about their real-time electricity prices and fuel diversity through websites or mobile apps. These services could allow savvy consumers to adjust their consumption in order to reduce their electricity bill or address environmental concerns. Information communication technologies (ICTs) could be such a consumer, shifting consumption in either time or space by changing the policies used to route data through the network.

Through energy-aware routing, it is possible to mitigate the electricity cost and emissions produced in core networks, but reducing both simultaneously can be difficult. For example, some types of conventional power plants (particularly coal power plants) are cheap sources of electricity, but they are also the world's top contributing sources of carbon dioxide emissions and the primary cause of global warming. Thus, reduction of electricity cost and reduction of carbon dioxide emissions may be distinct goals in opposition to one another. Previous works have typically focused on one goal, such as selecting the lowest electricity cost path between a source and destination in a network [4]. In this work, we seek to address both economic and environmental ICT concerns simultaneously through novel specialized energy-cost- and energy-source-aware routing

Manuscript received February 26, 2018; revised May 24, 2018; accepted May 29, 2018; published July 12, 2018 (Doc. ID 324673).

A. Deylamsalehi (e-mail: arash_deylamsalehi@student.uml.edu), D. A. P. Davis, P. Afsharlar, and V. M. Vokkarane are with the Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, Massachusetts 01854, USA.

M. Bahrami and W.-P. Chen are with Fujitsu Laboratories of America, Inc., Sunnyvale, California 94085, USA.

https://doi.org/10.1364/JOCN.10.000D72

1943-0620/18/100D72-12

Corrected 22 October 2018

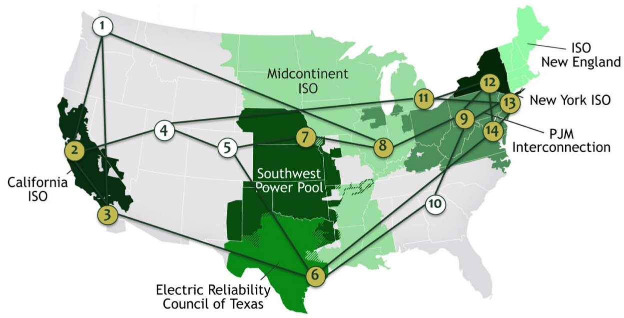

Fig. 1. 14-node NSFnet topology, with regional power markets highlighted.

paradigms as an interdisciplinary project between power grids and ICTs. Considering the electricity cost and emissions of power markets in a combined manner may yield superior economic green networking in optical networks to address both economic and environmental concerns about core networks as a joint concern.

Our approach comes in the form of a mixed-integer linear program (MILP) that adaptively assigns lightpath (a physical path and wavelength) solutions to the static routing and wavelength assignment (RWA) problem in optical core networks to optimally reduce either the electricity cost, emissions produced, or both. We examine minimizing the electricity cost and compare this objective to several different goals such as minimizing only the emissions, minimizing the lengths of provisioned routes, and finding a balanced emissions and dollar cost approach through a weighted cost function.

While the MILP can produce optimal solutions for a given objective, it does not scale well as the size of the network increases, as RWA is a NP-complete problem [5]. This can lead to high computation times, which can have negative repercussions, as prices and power sources fluctuate over time. To deal with this issue, we propose another dynamic routing algorithm, which enhances the network performance by employing the output of MILP as input for a machine learning solution. We study logistic regression for this purpose, and we compare its performance in terms of electricity cost and emissions accuracy.

In this paper, we build upon our previous work [6] by not only minimizing the electricity cost of RWA for a static set of requests, but also keeping emissions under a specified cap. This objective is quantitatively compared to additional alternative goals, including directly minimizing the emissions produced, reducing the length of established paths, and balancing reductions in both emissions and electricity cost simultaneously. The rest of this paper is organized as follows: we provide background information about electricity costs and emissions in power markets in Section III, describe our optical network nodal energy model in Section IV, define key equations in energy-aware routing in Section V, explain the MILP formulations in Section VI, investigate machine learning in Section VII and compare their performance in Section VIII, and conclude the paper in Section IX.

II. Related Work

A detailed survey of energy studies in the core networks has been presented by multiple works, including [7–9]. This subject has been considered from multiple perspectives, ranging from techniques for reducing the electricity cost of routing to cutting down on emissions. Given that ICT energy consumption is growing at ten times the average rate of all other sectors of electricity consumers [10,11], this is a critical area of study.

Electricity cost-aware routing was first introduced in Qureshi *et al.*'s paper in 2009 [12]. They distributed Internet traffic among different data centers based on their charged electricity prices. They ignored proportional power consumption in core networks used to transport that traffic and ignored the cost of using hubs located in regions without electricity markets. The authors of Ref. [13] similarly aimed to minimize electricity cost in multiple markets for Internet data-centers (IDCs). The same authors modeled the problem through mixed-integer programming based on a generalized benders decomposition (GBD) in Ref. [14]. In Refs. [15,16], joint optimization approaches were used that reduced both electricity and bandwidth cost in data-center networks. Cost savings was combined with fault tolerance in the design of data-center networks in Ref. [17]. Electricity cost minimization algorithms with job security guarantees for data-center networks in energy markets are presented in Ref. [18].

Power systems researchers have also investigated related topics. A mixed-integer programming model for participating data centers in demand-response programs was proposed by Ref. [19]. They evaluated their models using traffic data from the World Cup 1998 and found that more than 20% of electricity costs can be cut by taking the locational marginal electricity prices into account during the peak workload period. In Ref. [20], they simulated data centers interacting with the power grid with a flexible demand-response load. They proposed a pricing prediction model and examined the performance of power systems and flexible data centers during unstable traffic periods. A game-theoretic framework between a data center and its users was proposed by Ref. [21]. In this game, users can trade their performance for monetary rewards and shift their requests to different time periods. Electricity price contracts used by data centers in practice were reviewed in [22], and they found that data migration over WDM may reduce energy costs in data centers by up to 28%. The authors in Ref. [23] studied modular co-simulation of power systems and ICT to investigate how the performance of one impacts the cost of the other.

In Ref. [24], the goal was to minimize the total energy consumption in IP-over-WDM networks while minimizing its effects on the underlying core network's performance. In this work, they treated IP routers at the major sources of energy consumption. Similarly, in Ref. [25], mixed-integer linear programming (MILP) was used to cut electricity prices in an IP/WDM inter-data-center network with time-of-use pricing. The authors of Ref. [26] compared four optical core network architectures: (1) basic IP over WDM with no optical switching, (2) with transparent switching, (3) with translucent switching, and (4) IP over a synchronous digital hierarchy. They found savings of up to 60% in terms of energy costs by employing integer linear programming for energy minimization.

Beyond just considering the dollar cost of ICT, researchers have also examined the environmental impact. In Ref. [8], the green-energy-aware works are classified into four categories: green-energy-aware workload scheduling, green-energy-aware virtual machines (VM), green-energy-aware energy planning, and interdisciplinary. In Ref. [27], the authors suggest a new planning scheme to co-locate data centers and grid resources close to sites with a high availability of renewable resources. The trade-off between the rise in transport emissions and reduction of data-center emissions through this strategy has been investigated in Ref. [28]. Green-power-source-aware routing with generalized multiprotocol label switching is discussed in Ref. [29]. Renewable cloud services have been proposed to relocate traffic to cloud data centers with available renewable resources at the location [30]. In Ref. [31], the authors focused on reducing emissions by shifting energy consumption from non-renewable sources to solar energy. The authors of Ref. [32] present one of the first approaches that makes use of dual power sources for routing and wavelength assignment, and the authors of Ref. [33] investigate a heterogeneous, partially green network and its interaction with manycast routing. In Ref. [34], the GreenTouch consortium found that up to 98% reduction in the net energy consumption for the end-to-end communication networks can be achieved by 2020, through the use of a number of techniques. These include the installation of network components with lower power consumption, intelligent management of protection resources, using sleep mode and mixed line rates (MLR), optimization of the network physical topology, and using optimized content distribution throughout cloud networks [35]. When looking to mitigate the environmental impact at the core network scale, the proposed techniques reduce energy consumption to 1/39th its previous value for routers and to 1/6th for transponders.

Similarly, in Ref. [36], they performed green load balancing by routing the workload of cloud services to locations with a lower portion of "brown" fossil fuel energy. They found that emissions savings is inherently tied to the ability of a system to flexibly switch to lower-emissions power sources. The same authors studied the possibility of running a large-scale ICT systems using only renewable resources in Ref. [37]. The authors of Ref. [38] proposed a green routing algorithm, which worked by minimizing energy consumption in the core network, with the motivation that reducing energy consumption would have some effect on the amount of emissions produced.

As noted above, a number of emissions-reduction strategies for core networks have been proposed, but the concept of finding a balance between emissions and electricity cost in core networks has not yet been studied. To the best of our knowledge, there is no comprehensive analytical model that captures power consumption, emissions, and electricity cost for wide-area core networks. In this paper, we study the impact of adaptive routing based on emissions, electricity prices, and a combination of emissions and electricity cost.

III. Power Market Model

Power markets were created to foster competition in power generation and supply zones in terms of electricity price, reliability, quality, and even fuel diversity. They have administrators, like independent system operators (ISOs) and regional transmission organizations (RTOs), to make the market reliable, clear, and balanced. Due to lack of cost-effective technologies for storing electricity, the power markets are balanced by the principle of unit commitment, which ensures that the generated electricity must be equal to demand. This leads to fluctuation in electricity prices. There are seven main electricity markets in the United States (refer to Fig. 1), namely, New England ISO (NEISO), California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), Southwest Power Pool (SPP), Pennsylvania-Jersey-Maryland Interconnection (PJM), New York ISO (NYISO), and Midcontinent ISO (MISO). Figure 2 shows how the hourly electricity price changes at three different ISOs throughout a day [39].

Electricity is generated in power plants by using various origin resources; e.g., the United States fleet had a fuel diversity of coal (39%), natural gas (27%), nuclear (19%), hydro (6%), and other renewable resources (6%) in 2014 [40]. The process of generating electricity from each origin resource produces emissions at different rates, for instance: coal (800 to 1050 gCO_2e/kWh), natural gas (430 gCO_2e/kWh), nuclear (6 gCO_2e/kWh), hydro (4 gCO_2e/kWh), wind power (3 to 22 gCO_2e/kWh), and photovoltaic solar (60 to 150 gCO_2e/kWh) [41,42]. The gCO_2e is used as the emissions unit, and the "e" means all values include the impact of all kinds of emissions. Due to the lack of technologies for large-scale energy storage, there is no effective way to store large amounts of clean energy for later high-demand hours. Thus, fuel diversity across power markets and over time changes according to the cost and availability of origin resources. Based on fuel diversity and variety in the polluted air per kilowatt hour (kWh), we represent the total produced emissions in units of gCO_2e when consuming 1 kWh of electricity as the *emissions factor*. The emissions factor can fall between two extremes, ranging from 100% of the energy produced from a green fuel source to 100% produced by an environmentally damaging source. As we mentioned above, wind power plants can emit 3 gCO_2e per kWh (under ideal wind conditions), and are assumed to match the lowest emissions factor in this paper, while coal power plants match the highest rate at 1050 gCO_2e/kWh. Figure 3 shows emissions factors given at two ISOs.

IV. Power Consumption Model

There are different models for network equipment power consumption in the literature. We use a model that is based on real equipment measurements and accounts for power

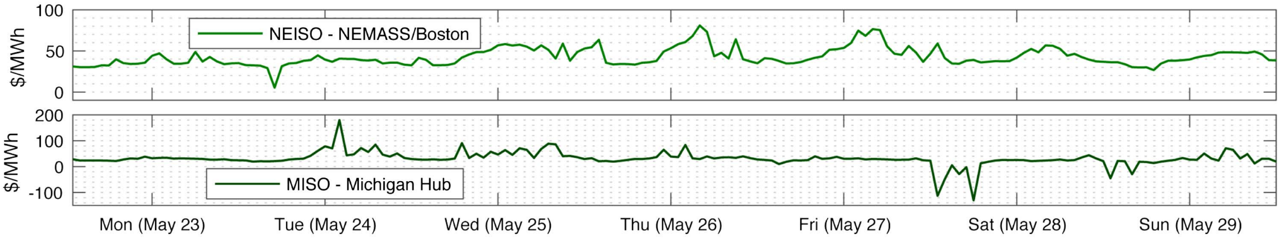

Fig. 2. Real-time hourly energy price for the week that started on May 23, 2011 at two different ISOs [39].

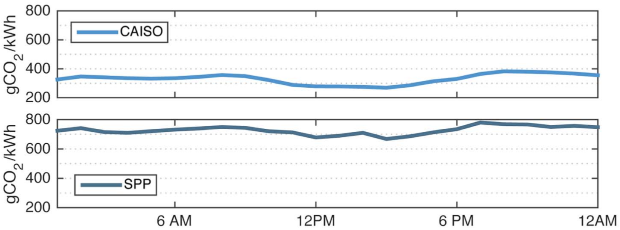

Fig. 3. Emissions factor for October 12, 2015 at two different ISOs.

consumption both while a node is on but idling and while the node is used for transmission. In Fig. 4 we illustrate the multilayer (optical, electronic, and IP layer) node model used to calculate the network power consumption. We use values from [43], where a comprehensive analysis of equipment from different manufacturers resulted in the quantities listed in Tables I and II. The node consists of an electrical routing layer at the top, an optical transport and multiplexing layer at the bottom, and a connecting optical electronic optical (OEO) layer in between. We use wavelength division multiplexing (WDM) as the transport network technology, as it is widely employed in high-speed wide area networks,

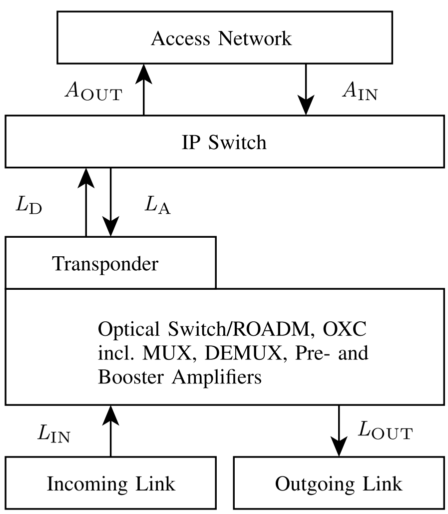

Fig. 4. Multilayer node power model.

$$p_{\mathrm{IP}} = \pi_{\mathrm{IP}}(A_{\mathrm{IN}} + A_{\mathrm{OUT}} + c(L_{\mathrm{D}} + L_{\mathrm{A}})), \tag{1}$$

$$p_{\mathrm{OEO}} = \pi_{\mathrm{TX}}(c)(L_{\mathrm{D}} + L_{\mathrm{A}}), \tag{2}$$

$$p_{\mathrm{WDM}} = \pi_{\mathrm{OXC}}(L_{\mathrm{D}} + L_{\mathrm{A}}) + \alpha(L_{\mathrm{IN}} + L_{\mathrm{OUT}}) + \beta, \tag{3}$$

$$P_j = p_{\mathrm{IP}} + p_{\mathrm{OEO}} + p_{\mathrm{WDM}}. \tag{4}$$

At the source, data are processed from the top, the access network, down to the outgoing link (refer Fig. 4), while data have to be processed in the reverse direction at destinations. Equations (1)–(4) are used to calculate the nodal power consumption along with the values in Tables I and II.

Equation (1) is the power consumption of the IP layer, and Eq. (2) describes the OEO power consumption. L_{IN}, L_D, and A_{OUT} are assumed equal to zero at sources, whereas A_{IN}, L_A, and L_{OUT} are zero at destinations. As only source and destination nodes involve OEO conversion, Eqs. (1) and (2) are not applied at any intermediate nodes along an all-optical route. Equation (3) considers the consumed electricity in optical communication, which is calculated at all involved nodes. α is equal to 0.085 kW and represents the power consumption of amplifiers built into the node. The β value

TABLE I
NETWORK-SPECIFIC VARIABLES

Symbol	Description	Value	Unit
c	Transponder transmission rate	10	Gbps
$\pi_{\mathrm{TX}}(c)$	Transponder power consumption	0.05	kW
π_{OXC}	OXC power consumption	0.1	kW
π_{IP}	IP layer power consumption	0.01	kW/Gbps

TABLE II
NODE-SPECIFIC VARIABLES

Symbol	Description	Unit
A_{IN}	Input traffic from access network	Gbps
A_{OUT}	Output traffic to access network	Gbps
L_{D}	Lightpaths dropped from WDM to IP (O-E)	Integer
L_{A}	Lightpaths added to WDM from IP (E-O)	Integer
L_{IN}	Incoming lightpaths	Integer
L_{OUT}	Outgoing lightpaths	Integer

is constant for directing the lightpaths to the appropriate ports and equals 0.15 kW. The total power consumption of a node is then calculated by summation of power consumption at all layers using Eq. (4).

V. Energy-Source-Aware Routing Equations

We now discuss how the output from the nodal power model can be utilized to calculate the energy consumption, electricity cost, and emissions produced from establishing a lightpath. A key element of this calculation is determining the dollar or emissions cost for each node when a request's lightpath traverses it.

First, the energy consumption of a node when used for a given request must be calculated. Following Section IV, let $\mathrm{serviceTime}_r$ be the duration of the data transfer for request r. Nodal energy consumption E_j^r at node j is found by multiplying the nodal power consumption at node j by $\mathrm{serviceTime}_r$,

$$E_j^r = P_j^r \times \mathrm{serviceTime}_r. \tag{5}$$

Therefore, given the electricity price at node j denoted by EP_j, the electricity dollar cost C_j^r of node j may be calculated as:

$$C_j^r = E_j^r \times EP_j. \tag{6}$$

As mentioned in Section III, the emissions factor γ_j at node j represents the production emissions in the scale of $\mathrm{gCO_2e}$ if 1 kWh energy is consumed at node j. Thus, the produced emissions Γ_j^r at node j for a given request may be calculated as:

$$\Gamma_j^r = E_j^r \times \gamma_j. \tag{7}$$

Using Γ_j^r or C_j^r as the weight for nodes, approaches in Section VI are able to select a lightpath for each request r so that the total emissions produced or dollars spent are minimized.

VI. Mixed-Integer Linear Programs

In this section, we introduce a MILP model that finds the optimal minimum dollar cost of provisioning a set of requests, incorporating the models and equations from Sections III–V. Following that, we propose several modifications for pursuing alternative objectives mirroring the routing methods outlined in Section V, which are minimizing the emissions produced and the number of hops per provisioned request, the physical distance traversed per request, and two additional objectives that balance emissions and dollar costs.

A. Least Dollar Paths

1) Given:

- V is the set of nodes in the network.
- W is the set of wavelengths available on each link.
- R is the set of unicast requests. For a given unicast request r, we denote the source node of the request as s_r, the destination node as d_r, and the amount of time the request needs to hold network resources as $\mathrm{serviceTime}_r$.
- $\zeta_{i,j}$ is binary, equal to 1 if a physical link exists between $i,j \in V$.
- $LEN_{i,j}$ is a non-integer, equal to the length in kilometers of link (i,j).
- EP_j is a non-integer equal to the dollar cost of a unit of energy at node j.
- γ_j is a non-integer equal to the emissions factor of energy spent at node j.
- P_j^r is a non-integer equal to the power consumption at node j when used for request r.
- E_j^r is the non-integer amount of power required to use node j for request r.
- C_j^r is the non-integer dollar cost of using node j for a request r. This cost is determined using Eq. (6).
- C_{MAX} is the maximum value of C_j^r across all combinations of $r \in R, j \in V$.
- Γ_j^r is the non-integer emissions cost of using node for request r. This cost is calculated using Eq. (7).
- Γ_{MAX} is the maximum value of Γ_j^r across all combinations of $r \in R, j \in V$.
- CAP_e is the non-integer emissions cap, which is a restriction on the amount of emissions produced from provisioning request set R.

2) Variables:

- $F_{i,j}^{r,w}$ is binary, equal to 1 when link (i,j) carries the lightpath for request r using wavelength w.
- $L^{r,w}$ is binary, equal to 1 when there is a lightpath for request r using wavelength w.
- N_j^r is binary, equal to 1 if the lightpath for request r goes through node j.
- DC^r is a non-integer, equal to the amount of dollars to establish the lightpath for request r.
- EC^r is a non-integer, equal to the amount of emissions produced when establishing the lightpath for request r.
- HC^r is an integer, equal to the number of hops in the provisioned path for request r.
- DIS^r is a non-integer, equal to the total length in kilometers of the provisioned path for request r.

3) Constraints: **Objective Function:**

$$\text{minimize:} \quad \sum_{r\in R} \mathrm{DC}^r. \tag{8}$$

The objective is to minimize the total dollar cost across all requests in the network [Eq. (8)].

Subject to:

$$\sum_{r}^{R} F_{i,j}^{r,w} <= 1; \quad \forall\, w \in W, \quad i,j \in V, \tag{9}$$

$$F_{i,j}^{r,w} \le \zeta_{i,j} \cdot L^{r,w}; \quad \forall\, r \in R, \quad w \in W, \quad i,j \in V, \tag{10}$$

$$\sum_{i}^{\overline{V}} F_{i,j}^{r,w} - \sum_{k}^{\overline{V}} F_{j,k}^{r,w} = \begin{cases} 0, & \text{if } j \neq s_r, d_r \\ L^{r,w}, & \text{if } j = d_r \\ -L^{r,w}, & \text{if } j = s_r \end{cases}$$
$$\forall j \in \overline{V}, \quad w \in \overline{W}, \quad r \in \overline{R}, \tag{11}$$

$$\sum_{w}^{W} L^{r,w} = 1; \quad \forall r \in R. \tag{12}$$

The constraints presented in Eqs. (9)–(12) are standard constraints for the unicast RWA problem. The constraint in Eq. (9) prevents a wavelength from being used more than once on a link, the one in Eq. (10) limits lightpaths to being established only on links that exist, and in Eq. (11) are the conventional flow conservation constraints. Equation (12) ensures that each request is provisioned a lightpath.

In the constraints presented in Eqs. (13) and (14), the variable N_j^r is set equal to 1 if there is at least one flow going into or departing from node j, and equal to 0 otherwise,

$$N_j^r <= \sum_{w}^{W} \sum_{i}^{V} F_{i,j}^{r,w} + \sum_{w}^{W} \sum_{k}^{V} F_{j,k}^{r,w}; \quad \forall r \in R, \quad j \in V, \tag{13}$$

$$N_j^r \cdot 2|W| >= \sum_{w}^{W} \sum_{i}^{V} F_{i,j}^{r,w} + \sum_{w}^{W} \sum_{k}^{V} F_{j,k}^{r,w}; \quad \forall r \in R, \quad j \in V. \tag{14}$$

Different cost variables are defined in the constraints presented in Eqs. (15)–(18), each assigning a type of cost to a request r. The constraint in Eq. (15) defines the dollar cost, in Eq. (16) the emissions cost, in Eq. (17) the path length in terms of hops, and in Eq. (18) the path length in terms of kilometers,

$$\mathrm{DC}^r = \sum_{j}^{V} C_j^r \cdot N_j^r; \quad \forall r \in R, \tag{15}$$

$$\mathrm{EC}^r = \sum_{j}^{V} \Gamma_j^r \cdot N_j^r; \quad \forall r \in R, \tag{16}$$

$$\mathrm{HC}^r = \sum_{w}^{W} \sum_{i}^{V} \sum_{j}^{V} F_{i,j}^{r,w}; \quad \forall r \in R, \tag{17}$$

$$\mathrm{DIS}^r = \sum_{w}^{W} \sum_{i}^{V} \sum_{j}^{V} F_{i,j}^{r,w} \cdot LEN_{i,j}; \quad \forall r \in R. \tag{18}$$

B. Least Emissions Paths

Rather than minimizing the total dollar cost [Eq. (8)], an alternative objective in Eq. (19) is to minimize the total emissions produced across all requests.

1) Constraints: **Objective Function:**

minimize:

$$\sum_{r}^{R} \mathrm{EC}^r. \tag{19}$$

C. Shortest Hop Paths

If delay is a concern, the total number of hops per request can be reduced by minimizing the sum across all requests using Eq. (20).

1) Constraints: **Objective Function:**

minimize:

$$\sum_{r}^{R} \mathrm{HC}^r. \tag{20}$$

D. Shortest Distance Paths

If the physical distance that an optical signal must travel is too great, signal regenerators may be necessary. The objective in Eq. (21) can be used to favor shorter distance paths.

1) Constraints: **Objective Function:**

minimize:

$$\sum_{r}^{R} \mathrm{DIS}^r. \tag{21}$$

E. Balanced Cost Paths

Rather than directly minimizing only the dollar cost or emissions, we can use a balanced approach that reduces both. As presented in Ref. [1], this approach uses a new weight function that will minimize the combined emissions and dollar cost based on weight value η ranging from 0.0 to 1.0. Note that the η is provided for a particular problem instance, and all requests must use the same η value.

Two new parameters (η and WC_j^r), a new variable (BC^r), a new objective [Eq. (23)], and an additional constraint [Eq. (24)] are needed to minimize this balanced weighted cost.

1) Given:

- η is a non-integer weight value, falling within the range [0.0, 1.0] in increments of 0.1, i.e., 0.0, 0.1, 0.2,…,0.9, 1.0.
- WC_j^r is non-integer, equal to the weighted combined cost of emissions and dollars for using node j in the lightpath for request r. It is calculated using Eq. (22).

2) Parameter Equation:

$$\mathrm{WC}_j^r = \eta \cdot \frac{\Gamma_j^r}{\Gamma_{\mathrm{MAX}}} + (1-\eta) \cdot \frac{C_j^r}{C_{\mathrm{MAX}}}; \quad \forall r \in R, \quad j \in V. \tag{22}$$

3) Variables:

- BC^r is a non-integer equal to the total weighted cost of a lightpath for request r.

4) Constraints: **Objective Function:**

minimize:

$$\sum_{r}^{R} \mathrm{BC}^r. \tag{23}$$

Subject to:

$$\mathrm{BC}^r = \sum_{j}^{V} \sum_{w}^{W} WC_j^r \cdot N_j^r; \quad \forall\, r \in R. \tag{24}$$

F. Economic Green Paths

The weighted cost described in the *Balanced Cost Paths* subsection uses the same weight value η for every request in a request set. To improve the flexibility of the balanced cost approach, we present the new method, economic green paths (EGP), in this subsection. An entire set of weight values $H = \{0.0, 0.1, 0.2, \ldots, 0.9, 1.0\}$ are made available for EGP, and we alter the definition of the WC_j^r parameter to use whichever weight value from H will minimize the balanced cost for each individual request. This opens up the potential for greater emissions or cost savings on a case-by-case basis.

1) Given:

H is a set of non-integer weight values, with each $\eta \in H$ ranging from 0.0 to 1.0 in increments of 0.1.

The focus for this formulation is shifted to reducing the total dollar cost [Eq. (25)], while keeping the total emissions under a provided emissions cap [Eq. (26)].

Objective Function:

minimize:

$$\sum_{r \in R} \mathrm{DC}^r. \tag{25}$$

Subject to:

$$\sum_{r}^{R} \mathrm{EC}^r <= \mathrm{CAP}_e. \tag{26}$$

CAP_e is chosen based on network operator policies. This cap does open up the potential for creating infeasible scenarios, where it is possible that the MILP will not be able to produce a solution if the CAP_e is set too low.

In this work, we defined the emissions cap based on the percentage of emission reduction over LDP. By using the equations in Ref. [1], the emissions cap may calculated by:

$$\mathrm{CAP}^e = \Gamma_{\mathrm{LDP}} - \mathrm{reductionRate} \cdot (\Gamma_{\mathrm{LDP}} - \Gamma_{\mathrm{LEP}}), \tag{27}$$

where CAP^e is the emission cap determined by the reduction rate, along with Γ_{LDP} and Γ_{LEP}, which are the emissions produced by LDP and LEP, respectively. This implies that the lowest cap is equal to the emissions produced by using LEP routing, while the highest cap is equal to the emissions from LDP routing.

VII. Machine Learning

MILP, while producing the optimal solution for a given objective, is not scalable. As the size of the network or the number of requests increases, the solving time grows exponentially. As prices are known to fluctuate throughout a day, operators would be interested in finding a solution that can give them effective savings in terms or dollar cost or emissions within a reasonable time frame. We present a technique that employs a well-known supervised machine learning algorithm, logistic regression, that predicts η in real time after it has been provided a training set of optimal MILP solutions. This can be used by a network operator to choose a cost-efficient route, given a set of options and the dollar/emissions costs for those routes, in a much shorter amount of time than dynamically calculating the solution using MILP.

In this section, we explain how we extract the key points of the input for training a logistic regression model (feature extraction on MILP), the implementation of the model, as well as its output results and the accuracy of the predictions of the model. The trained model can be used to compute η in real time. We use a set of features, including a set of dynamic network configurations, to train the model. Our goal is to train the model with a varied set of network configurations so that the model is able to make predictions when presented with configurations it has not been previously trained on. Since the trained model relies on computed statistical results, and it does not rely on network configuration, it allows a network provider to predict η in real time.

A. Binary Value for η

In terms of a single request, changing the η from zero to one changes the weight factor of potential paths between a source and a destination. When $\eta = 0$, the link weights only include the emissions cost of transmission. When $\eta = 1$, the weight considers only the dollar cost. As η increases from zero to one, there is a tipping point η_i where the dollar cost will have a heavier influence on the weight than the emissions. In BCP, any provided η less than η_i will result in the selection of the minimum emissions path, while any η greater than η_i will force BCP to choose the minimum dollar cost path. Note that any path between a source and a destination has a potential electricity cost and emissions value, which is completely independent of η, as it is based on the nodal electricity price and emissions factor for each node in the path.

In contrast to BCP where η is supplied as a parameter, EGP picks the required η for any request while simultaneously considering the associated electricity cost and emissions on that path. As discussed above, a number of η values within certain ranges may lead to the same decision on which path to take. In fact, based on our observations, more than 99.7% of requests end up choosing an η value of either zero or one. This indicates that the overall optimal weighted cost can be found by choosing either the minimum emissions path or the minimum dollar cost path, rather than choosing a balanced path that falls between those two extremes. Keeping this in mind, we simplify the decision made by the logistic regression model covered in the following subsection by limiting its choice in η to either zero or one.

B. Feature Extraction

The first step to generating a machine learning model is feature extraction from a pre-computed η that allows a model to be trained with different types of input. The goal of this extraction is to allow the model to understand different future inputs and predict the final value of η. This avoids the need for η computation using MILP for new requests, which would be costly.

While we extracted a number of features from the input, the features covered in this section make up the most well-defined model that provides a highly accurate prediction rate.

We consider several binary dependent variables for training a logistic regression to predict η value from the MILP. The input data is made up of a set of requests that can be optimized. This group of requests can end up taking one of a number of paths by changing η. In this work, we called this the "Flexible Request Set." This set is made up only of the requests that end up having a number of routing choices based on their η value. This comes out to 20% of our total set of requests, with the remaining 80% selecting the same path for both LDP and LEP objectives. MILP provides the following information for training the machine learning module for each request in the flexible request set:

- CAP_e We set the emission cap for a week's worth of requests in MILP. The emissions of traffic outside of the flexible request set are subtracted from the total emissions cap. This value is identical for all requests during the week with same emissions reduction rate.
- s Source node.
- d Destination node.
- t_h Holding time.
- $\mathrm{EC}_{\mathrm{LDP}}$ Total emissions produced by path selected by LDP.
- $\mathrm{EC}_{\mathrm{LEP}}$ Total emissions produced by path selected by LEP.
- $\mathrm{EC}_{\mathrm{LDP}}$ Total electricity cost for path selected by LDP.
- $\mathrm{EC}_{\mathrm{LEP}}$ Total electricity cost for path selected by LEP.
- η : Selected η by EGP MILP.

C. Logistic Regression Model

We used extracted features to train a logistic regression model, which is a linear model for the classification problem of predicting a value η for a given pre-computed, s, d, t_h, $\mathrm{EC}_{\mathrm{LDP}}$, $\mathrm{EC}_{\mathrm{LEP}}$, $\mathrm{EC}_{\mathrm{LDP}}$, $\mathrm{EC}_{\mathrm{LEP}}$. The model can be formulated as follows. There are two types of regularization as loss functions for logistic regression: i) L1-norm and ii) L2-norm loss function. L1 regularized logistic regression solves the following optimization problem:

$$\min_{w,c} \|w\|_1 w^T + C \sum_{i=1}^{n} (\log(\exp(-y_i) X_i^T w + c)) + 1, \tag{28}$$

$$\min_{w,c} \frac{1}{2} w^T + C \sum_{i=1}^{n} (\log(\exp(-y_i) X_i^T w + c)) + 1. \tag{29}$$

Since we have a large dataset of computed variables, we can use a solver with L2 penalization that also converges faster for high dimensional data. The solver uses stochastic average gradient descent [44], which is an objective function as follows:

$$Q(w) = \frac{1}{n} \sum_{i=1}^{n} Q_i(w). \tag{30}$$

The iteration model of Eq. (30) is defined as follows:

$$w \coloneqq w - \eta \nabla Q(w) = w - \eta \sum_{i=1}^{n} \nabla Q_i(w)/n. \tag{31}$$

VIII. Numerical Evaluation

In this section, we compare LDP, LEP, BCP, SHP, SDP, and EGP objectives in terms of emissions and electrical expenditures. We evaluate MILP performance for these objective using the 14-node NSFnet topology shown Fig. 1. We follow this up by examining the accuracy of our logistic regression model in predicting the η value given a training set made up of MILP solutions.

A. MILP Evaluation

The MILP solutions considered use adaptive routing, where a route is chosen based on the current network state, which can provide opportunities for savings by choosing intermediate nodes based on the price and emissions factor currently tied to those nodes. This improvement is most noticeable when there are differences between the electricity prices or emissions factors across the nodes in the network. If real-time pricing is available, the magnitude of these differences changes as a day progresses, for a variety of reasons. For example, costs may be lower outside of peak usage hours, and different regions of a network may experience peak traffic at different times due to time-zone differences.

We consider a number of Emission cap scenarios for EGP, ranging from 0% reduction from LDP, to 25%, 50%, 75%, and 100% reduction (identical to LEP emissions performance). These caps can change the behavior of EGP, as dollar cost savings may have to be compromised to stay under the emissions cap.

We average our results over 30 seeds, with each set of requests made up of 16,560 requests to uniform randomly chosen sources and destinations, and average holding times of 1 h.

Requests are spread over 168 h (a week), with different emissions and power cost values based on the time period that they will be active. The distribution is based on the hourly traffic model in Ref. [27], where network traffic peaks in the evening and falls during the early morning hours. The location of the source node in respect to different time zones is taken into account. Each request begins its holding time at the start of the hour in which they arrive to the network. Requests are assumed to be comprised of multiple traffic flows, consolidated from the electronic layer, such that each established lightpath consumes approximately the capacity of one wavelength. The number of wavelengths available on a length, $|W|$, is set to 12 on each link.

Each seed is solved using the Gurobi linear program solver [45], run as a job on the Massachusetts High Performance Green Computing Center [46], and allocated twelve cores. Each seed was completed within a 24 h period.

Figure 5 compares the weekly totals for emissions produced and dollars spent for LDP, LEP, and BCP MILP objectives. Lines are included to show the values used for emissions caps in later graphs. Immediately noticeable is the curve ranging from LEP to LDP, with η weight values in between ranging from 0.9 to 0.1, each representing a weight used for the BCP approach. Each η value corresponds to an average value for BCP given that η value. LEP, which is equivalent to BCP with $\eta = 0$, consumes the fewest emissions (approximately 13740 $kgCO_2e$), LDP (BCP with $\eta = 1.0$) the fewest dollars (approximately \$1565), and BCP covers the curve in between. This graph shows the trade-off that must be made when minimizing dollar cost or emissions: focusing on reducing dollars consequently increases the emissions, and attempting to minimize emissions produced limits on the dollar savings possible. The cheapest nodes in terms of dollar cost may often not be the cheapest in terms of emissions during any given hour, so a choice must be made between them when routing from a source to a destination.

Figure 6 compares the weekly totals for emissions produced and dollars spent for each LDP, LEP, shortest hop path (SHP), shortest distance path (SDP), BCP, and EGP. SHP, being dollarcost blind, costs more than any of the energy-focused approaches, but approximately matches the BCP approach with a weight of 0.7 in terms of emissions. This is compared to SDP, which chooses paths based on minimizing the physical distance that a path traverses

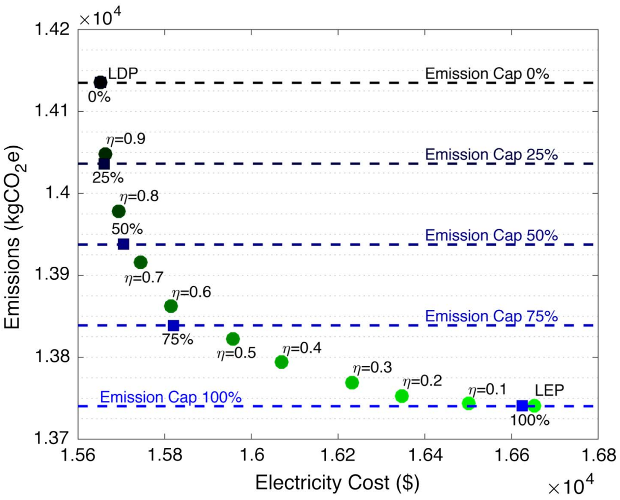

Fig. 5. Total electricity expenditure versus total emissions for LDP, LEP, and BCP with various η during the week.

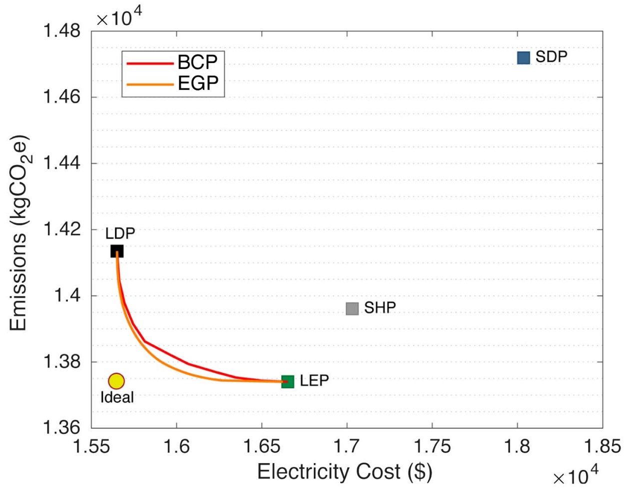

Fig. 6. Total electricity expenditure versus total emissions for LDP, LEP, BCP, SDP, SHP, and EGP with various η during the week.

and performs by far the worst in regards to both emissions and dollars. This highlights a key takeaway of the evaluation. The total dollar cost and emissions is dependent on how many nodes are included in the path, as costs are incurred for each node used for transmission. The physical length of each network link has no correlation with the dollar or emissions cost of using a node, so focusing on reducing path length provides no savings. Note that we have also included an ideal point on the graph where both the total emissions (LEP) and the total dollar cost (LDP) are minimized, for comparison with our other objectives.

Over the course of the simulated week, real-time electricity price decreases after midnight, whereas in regions where a flat rate is used the price remains the same. LDP, BCP, and EGP, as they factor in the dollar cost of the route, can avoid transferring data through the flat-rate nodes during those cost-saving periods as much as possible. The emissions produced in different regions can vary at different times as well. Weather conditions may change the performance of renewable resources. Sunny or windy weather can increase the share of renewable resources in the market through solar or wind power sources and lead to lowering the emissions factor for nodes that use those resources. LEP, BCP, and EGP therefore tend to pass the data through the path with nodes that utilize lower-emissions power sources.

As an overall conclusion to our evaluation, the differences between the LEP, LDP, BCP, and EGP solutions show the choices network operators will have to consider in trying to reduce cost (either dollars or emissions) in optical core routing. If dollars are the only concern, LDP can be used for a routing policy, while LEP is appropriate for minimizing emissions. BCP can provide solutions (particularly around $\eta = 0.7$) that can provide both moderate cost savings and a noticeable reduction of impact on the environment. EGP can go even further by making a decision for each request, rather than the entire request set, on

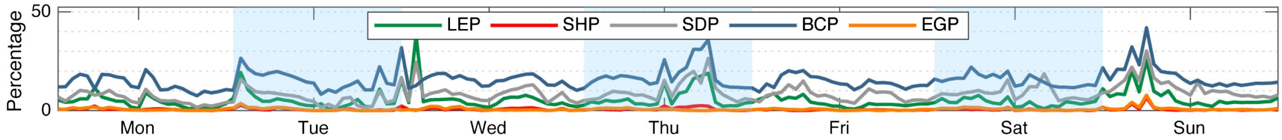

Fig. 7. Electricity cost comparison of LEP, BCP, SHP, SDP, and EGP versus LDP in 14-node NSFnet.

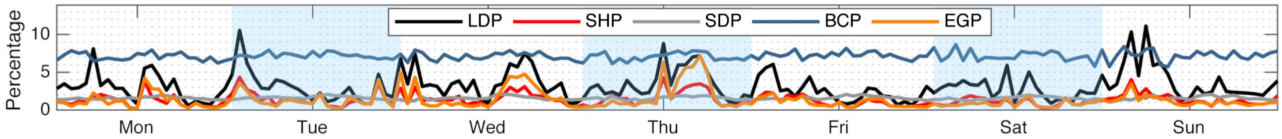

Fig. 8. Emissions comparison of LDP, BCP, SHP, SDP, and EGP versus LEP in 14-node NSFnet.

whether it is more effective to prioritize reducing dollar cost or emissions. This case-by-case flexibility enables it to avoid selecting lengthy paths that, while potentially cheap in terms of dollars or emissions, can be very costly in terms of the other metric. The electricity cost of the BCP ($\eta = 0.7$), SHP, SDP, and EGP (with a 50% emissions cap) objectives are compared to the optimal LDP in Fig. 7. A similar comparison to LEP in terms of emissions is shown in Fig. 8. These two graphs illustrate how the costs and emissions can fluctuate over time, both within a single day and over the whole week.

B. Evaluation of Logistic Regression Model

We used a well-known machine learning library [47] for implantation of a logistic regression with a L2 penalty in Python. The input variable consists of 10 features that predict η for the Section VI.F MILP. The dataset is divided into two parts, with some percentage for training and the remaining for testing the accuracy of the model's predictions.

Figure 9 shows the results of the logistic regression evaluation of the model for different training datasets with different sizes. The logistic regression model aims to predict the η value in order to offer a real-time alternative to MILP for solving this problem.

Figure 9 shows the accuracy of the predictions, and the predicted η values themselves, given different partitions of the dataset into training and testing subsets. For example, 10% testing size indicates that 90% of the data have been used for training the model and 10% of records from the dataset have been used for testing the model. The testing verifies the accuracy of the model for predicting unseen records in the training dataset. The figure also shows an accuracy bar for each selected dataset size and η value line for different selected training dataset sizes. As the number of requests in the training set is increased, the predicted η comes closer to the actual average η Fig. 9. The model is most accurate when the dataset is made up of 808,024 and 738,024 records.

The best results show 92.5% accuracy for predicting η, which happens when 70% of the data is used for training and 30% for testing purposes, which is 42.5% better than random selection. By employing machine learning with 92.5% accuracy, given that a number of requests have been calculated ahead of time and used for training, the operator can often find a cost-effective path between any source and destination by using the η output of the model to determine which path to choose. Although the best model, with 70% and 30% of records for training and testing, respectively, shows the best performance, even smaller training and testing datasets show reasonable and efficient results compared to the random baseline. For instance, the difference in accuracy between the largest dataset (808,024 records) and the smallest dataset (608,024 records) is less than 10%. This means that even when considering only the minimum

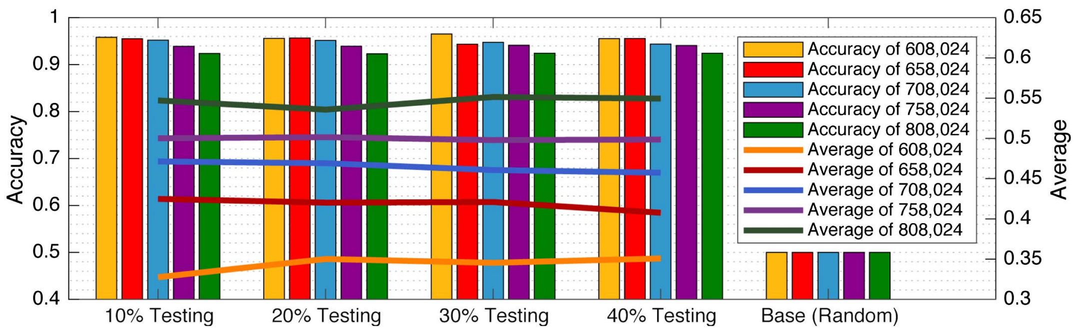

Fig. 9. Logistic regression model results.

number of records, we are still able to achieve more than 80% accuracy when predicting the value of η. We also examine our proposed solution for the larger 24-node USnet [1] and similarly compare the machine learning results to the MILP results. The results show 92.8% accuracy for predicting η, which happens when 70% of the data is used for training and 30% for testing purposes. Given the similarity between the results for NSFnet and USnet, we have omitted this additional data due to space constraints.

IX. Conclusion

Economic and environmental restrictions are two main criteria for any energy consumer. Neither can be improved substantially without impacting the other. In this paper, we investigated approaches for mitigating the environmental and economic concerns of ICT simultaneously by employing MILP and machine learning. Our approach promises economical and ecological improvement for data-center network operations by formulating electricity cost and emissions of optical data-center networks as a multi-objective function that can balance between minimizing the electricity cost and the emissions. The numerical results show that the proposed approach can be an effective solution for network operators looking to improve the dollar or emissions efficiency of their network. Through employing a logistic regression model, we found an effective way for network operators to choose an economic green path with little computation time, given that the model has been sufficiently trained. This will enable operators to react to changes in price and emissions in real time, maximizing the potential for savings for both themselves and the environment. The model is able to predict the η value with accuracy of up to 92.5% by considering a dataset of 808,024 records when it is divided into 70% and 30% for training and testing, respectively.

This topic is open for future investigation. Additional topologies and pricing models can be explored to determine if our conclusions hold under different conditions. This paper proposes basic tools for participation of ICT in the power markets. There are a number of areas for future work to explore. Our model does simplify some factors, such as the power consumption of regenerators, which are necessary for transmitting over long distances. Future work can incorporate such factors to add more realism to the model. A sizable percentage of data-center network traffic is machine-to-machine, which is mostly flexible with regard to scheduling. Further investigation can lead to the development of new algorithms to shift traffic in the time domain and merge this shift in time with flexible selection of geographical locations, which can help maximize the benefit of participation in the demand response program. Paradigms such as advance reservation or delayed allocation could incorporate load shifting when scheduling traffic. On the other hand, big electricity consumers buy the granted part of their coming day consumption from day-ahead markets then provide the rest of their needed electricity from real-time markets. More investigation is required to analyze the combination of these two markets for optical data-center networks and to propose a novel manner to predict the network traffic needs for the day ahead. Network operators can shift the load in location or time, changing the amount of energy drawn from either of these two markets to achieve greater savings.

Acknowledgment

This work has been supported by the NSF CARGONET project under grant CNS-1406370.

References

[1] A. Deylamsalehi, Y. Cui, P. Afsharlar, and V. M. Vokkarane, “Minimizing electricity cost and emissions in optical data center networks,” *J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. 257–274, 2017.

[2] B. Alcott, “Historical overview of the Jevons paradox in the literature,” in *The Jevons Paradox and the Myth of Resource Efficiency Improvements*, 2008, vol. 8.

[3] A. Qureshi, “Power-demand routing in massive geo-distributed systems,” Ph.D. dissertation, Massachusetts Institute of Technology, 2010.

[4] A. Deylamsalehi, T. Schondienst, and V. M. Vokkarane, “Real-time energy price aware network routing,” in *11th Annual IEEE High-capacity Optical Networks and Emerging/Enabling Technologies (HONET)*, 2014, pp. 15–19.

[5] I. Chlamtac, A. Ganz, and G. Karmi, “Lightpath communications: an approach to high bandwidth optical WAN’s,” *IEEE Trans. Commun.*, vol. 40, no. 7, pp. 1171–1182, July 1992.

[6] A. Deylamsalehi, D. A. Davis, and V. M. Vokkarane, “Electricity cost and emissions reduction in optical networks,” in *Int. Conf. on Computing, Networking and Communications (ICNC)*, 2017, pp. 381–386.

[7] A. P. Bianzino, C. Chaudet, D. Rossi, and J.-L. Rougier, “ A survey of green networking research,” *IEEE Commun. Surv. Tutorials*, vol. 14, no. 1, pp. 3–20, 2012.

[8] F. Kong and X. Liu, “A survey on green-energy-aware power management for datacenters,” *ACM Comput. Surv.*, vol. 47, no. 2, p. 1–38, 2015.

[9] F. Idzikowski, L. Chiaraviglio, A. Cianfrani, J. L. Vizcaino, M. Polverini, and Y. Ye, “A survey on energy-aware design and operation of core networks,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1453–1499, 2016.

[10] J. Koomey, “Growth in data center electricity use 2005 to 2010,” A report by Analytical Press, completed at the request of The New York Times, vol. 9, 2011.

[11] G. Ghatikar, “Demand response opportunities and enabling technologies for data centers: findings from field studies,” LBNL-5763E, 2014.

[12] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, “Cutting the electric bill for Internet-scale systems,” *ACM SIGCOMM Comp. Commun. Rev.*, vol. 39, no. 4, pp. 123–134, 2009.

[13] L. Rao, X. Liu, L. Xie, and W. Liu, “Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment,” in *Proc. IEEE INFOCOM*, 2010, pp. 1–9.

[14] L. Rao, X. Liu, M. Ilic, and J. Liu, “MEC-IDC: joint load balancing and power control for distributed internet data centers,” in *Proc., 1st ACM/IEEE Int. Conf. on Cyber-Physical Systems*, 2010, pp. 188–197.

[15] Y. Li, H. Wang, J. Dong, J. Li, and S. Cheng, "Operating cost reduction for distributed internet data centers," in *13th IEEE/ACM Int. Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, IEEE, 2013, pp. 589–596.

[16] H. Xu and B. Li, "Cost efficient datacenter selection for cloud services," in *1st IEEE Int. Conf. on Communications in China (ICCC)*, 2012, pp. 51–56.

[17] R. Tripathi, S. Vignesh, and V. Tamarapalli, "Cost-aware capacity provisioning for fault-tolerant geo-distributed data centers," in *8th Int. Conf. on Communication Systems and Networks (COMSNETS)*, IEEE, 2016, no. 1–8.

[18] Z. Li, J. Ge, C. Li, H. Yang, H. Hu, B. Luo, and V. Chang, "Energy cost minimization with job security guarantee in internet data center," *Future Gener. Comput. Syst.*, vol. 73, pp. 63–78, 2017.

[19] J. Li, Z. Li, K. Ren, and X. Liu, "Towards optimal electric demand management for Internet data centers," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 183–192, 2012.

[20] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 42, no. 1, pp. 111–123, 2014.

[21] Y. Zhan, D. Xu, H. Yu, and S. Yu, "Incentivizing users of data centers participate in the demand response programs via time-varying monetary rewards," arXiv:1604.01950, 2016.

[22] B. Aksanli, J. Venkatesh, T. Rosing, and I. Monga, "A comprehensive approach to reduce the energy cost of network of datacenters," in *IEEE Symposium on Computers and Communications (ISCC)*, July 2013, pp. 275–280.

[23] S. C. Müller, H. Georg, C. Rehtanz, and C. Wietfeld, "Hybrid simulation of power systems and ict for real-time applications," in *3rd IEEE PES Int. Conf. and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*, 2012, pp. 1–7.

[24] P. Saengudomlert, "Power-aware logical topology optimization for IP-over-WDM networks based on per-lightpath power consumption model," *ECTI Trans. Comput. Inf. Technol.*, vol. 9, no. 1, pp. 46–54, 2015.

[25] B. Kantarci and H. T. Mouftah, "Inter-data center network dimensioning under time-of-use pricing," *IEEE Trans. Cloud Comput.*, vol. 4, no. 4, pp. 402–414, 2016.

[26] F. Musumeci, M. Tornatore, and A. Pattavina, "A power consumption analysis for IP-over-WDM core network architectures," *J. Opt. Commun. Netw.*, vol. 4, no. 2, pp. 108–117, 2012.

[27] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Commun. Surv. Tutorials*, vol. 12, no. 4, pp. 441–458, 2010.

[28] M. Gattulli, M. Tornatore, R. Fiandra, and A. Pattavina, "Low-carbon routing algorithms for cloud computing services in IP-over-WDM networks," in *IEEE Int. Conf. on Communications (ICC)*, 2012, pp. 2999–3003.

[29] J. Wang, S. Ruepp, A. Manolova, L. Dittmann, S. Ricciardi, and D. Careglio, "Green-aware routing in GMPLS networks," in *Int. Conf. on Computing, Networking and Communications (ICNC)*, 2012, pp. 227–231.

[30] U. Mandal, M. Habib, S. Zhang, B. Mukherjee, and M. Tornatore, "Greening the cloud using renewable-energy-aware service migration," *IEEE Netw.*, vol. 27, no. 6, pp. 36–43, 2013.

[31] X. Dong, T. El-Gorashi, and J. M. Elmirghani, "IP over WDM networks employing renewable energy sources," *J. Lightwave Technol.*, vol. 29, no. 1, pp. 3–14, 2011.

[32] S. Ricciardi, F. Palmieri, U. Fiore, D. Careglio, G. Santos-Boada, and J. Solé-Pareta, "An energy-aware dynamic RWA framework for next-generation wavelength-routed networks," *Comput. Netw.*, vol. 56, no. 10, pp. 2420–2442, 2012.

[33] T. Schöndienst, D. A. Davis, J. M. Plante, and V. M. Vokkarane, "Renewable energy-aware manycast overlays," *IEEE J. Select Areas Commun.*, vol. 32, no. 8, pp. 1585–1599, 2014.

[34] GreenTech, "*Greentouch final results from green meter research study: reducing the net energy consumption in communications networks by up to 98,*" White Paper, https://s3-us-west-2.amazonaws.com/belllabs-microsite-greentouch/uploads/documents/GreenTouch_Green_Meter_Final_Results_18_June_2015.pdf.

[35] J. Elmirghani, T. Klein, K. Hinton, L. Nonde, A. Lawey, T. El-Gorashi, M. Musa, and X. Dong, "Greentouch greenmeter core network energy-efficiency improvement measures and optimization," *J. Opt. Commun. Netw.*, vol. 10, no. 2, pp. A250–A269, 2018.

[36] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew, "Greening geographical load balancing," in *Proc. ACM SIGMETRICS Joint Int. Conf. on Measurement and Modeling of Computer Systems*, 2011, pp. 233–244.

[37] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew, "Geographical load balancing with renewables," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 39, no. 3, pp. 62–66, 2011.

[38] A. P. Bianzino, C. Chaudet, F. Larroca, D. Rossi, and J.-L. Rougier, "Energy-aware routing: a reality check," in *IEEE GLOBECOM Workshops (GC Wkshps)*, 2010, pp. 1422–1427.

[39] https://www.ferc.gov/market-oversight/mkt-electric/overview.asp.

[40] http://www.eia.gov/tools/faqs/faq.cfm?id=427&t=3.

[41] http://www.eia.gov/tools/faqs/faq.cfm?id=74&t=11.

[42] http://www.manicore.com/anglais/missions_a/carbon_inventory.html.

[43] W. Van Heddeghem, F. Idzikowski, W. Vereecken, D. Colle, M. Pickavet, and P. Demeester, "Power consumption modeling in optical multilayer networks," *Photon. Netw. Commun.*, vol. 24, no. 2, pp. 86–102, 2012.

[44] M. Schmidt, N. Le Roux, and F. Bach, "Minimizing finite sums with the stochastic average gradient," *Math. Program.*, vol. 162, no. 1–2, pp. 83–112, 2017.

[45] Gurobi Optimization, "The state-of-the-art mathematical programming solver," [Online]. Available: http://www.gurobi.com/.

[46] Massachusetts Green High Performance Computing Center [Online]. Available: https://mghpcc.org.

[47] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.