

Technical Report – Renewable Energy-Aware Information-Centric Networking

Julien Mineraud, Liang Wang, Sasitharan Balasubramaniam, *Senior Member, IEEE*
and Jussi Kangasharju, *Member, IEEE*

Abstract—The ICT industry today is placed as one of the major consumers of energy, where recent reports have also shown that the industry is a major contributor to global carbon emissions. While renewable energy-aware data centers have been proposed, these solutions have certain limitations. The primary limitation is due to the design of data centers which focus on large-size facilities located in selected locations. This paper addresses this problem, by utilizing in-network caching with each router having storage and being powered by renewable energy sources (wind and solar). Besides placing contents closer to end users, utilizing in-network caching could potentially increase probability of capturing renewable energy in diverse geographical locations. Our proposed solution is dual-layered: on the first layer a distributed gradient-based routing protocol is used to discover the paths along routers that are powered by the highest renewable energy, and on the second layer, a caching mechanism will pull the contents from the data centre and place them on routers of the paths that are discovered by our routing protocol. Through our experiments on a testbed utilizing real meteorological data, our proposed solution has demonstrated increased quantity of renewable energy consumption, while reducing the workload on the data centers.

Index Terms—renewable energy, energy-aware routing, in-network caching

1 INTRODUCTION

THE current global energy depletion is a widely debated topic, due largely to the increasing population which has fueled massive industrial growth in the last few decades. Parallel to the pressing problem of energy depletion, is the increase in CO_2 emissions [?] which has affected the environment in the form of global warming. These problems have largely been attributed to the poor planning process and short term goals that we as humans have taken in utilizing natural resources, over the years. We are now witnessing a closer relationship between ICT and its influence in the energy sector [?], [?]. In particular, the growth of the Internet is slowly moving up the ranks as a major source for energy consumption (10% of the world global energy consumption [?]), which is close to other established industries (e.g. airline industry). The wide spread popularity of the Internet has led to an increasing number of deployed communication networks (e.g. WiFi, WiMAX) providing rich services (e.g. Multimedia contents) to end users' devices. These rich services are usually represented through contents that are placed in high powered data centers [?], which today is one of the major source of energy consumption in the whole Internet infrastructure.

This new landscape has shifted ICT researchers towards developing solutions that can improve energy consumption of communication networks, and at the

same time minimize CO_2 emissions. In particular, as we witness increasing developments in renewable energy infrastructure, ICT researchers are pursuing new solutions where clean energy could be used as an energy source for the Internet infrastructure (e.g. designing energy efficient networks). While new solutions have been proposed for increasing renewable energy sources for data centre networks, the limitation of these solutions are the fact that the number of data centers are only localized in small number of locations, which also minimizes the probability of meeting high quantity of renewable sources (e.g. as we know, renewable energy is highly dependent on the weather patterns as well as location). However, compressing the size of data centers and increasing their distributed locations, will require a brand new design, which will incur high infrastructure costs. Therefore, a more feasible and practical solution is required for distributing the source of content storage, that is flexible, highly dynamic and reactive to the changing weather patterns.

In this paper, we propose a solution that meets this objective, by utilizing content routers as a distributed source for content storage. Storing contents on the routers provides a number of appealing benefits. Firstly, storing contents within the network provides an opportunity to bring the contents closer to the clients, which minimizes the need to re-fetch the contents from end data centers. Secondly, the lower costs of storage prices, means that changes in the infrastructure costs would be minimal in order to enable content caching within the network. However, the main benefit that also suits our proposed solution, is the fact that contents can be stored in a distributed manner and increases the distribution of

• J. Mineraud, L. Wang and J. Kangasharju are with the Department of Computer Science, University of Helsinki, Finland.
E-mail: firstname.lastname@cs.helsinki.fi.

• S. Balasubramaniam is with the Department of Electronic and Communication Engineering, Tampere University of Technology, Tampere, Finland.
E-mail: sasi.bala@tut.fi

locations and access to wider renewable energy sources (we assume that each router is powered by renewable energy infrastructure, which could be a combination of wind turbines and solar panels). Therefore, by pulling contents out from the data centers and placing them on content routers which are powered by renewable energy, we will have an opportunity of accessing higher quantity of renewable energy to power the content, which in turn can allow us to power-off certain servers in the data centers that may utilize brown energy.

The novelty of our work lies in the ability to maximize the use of renewable energy to power storage points for contents. Our solution is dual-layered:

- The first layer uses a gradient-based routing algorithm to discover paths along the routers that have access to high renewable energy.
- The second layer uses a CCN-like [?] in-network caching to cache the contents along the discovered paths.

A thorough evaluation has been conducted using real renewable energy data, on a real network topology. The evaluation was validated using an experimental testbed. Our key findings can be summarized as follows:

- The combination of our routing and caching is very effective at both increasing use of renewable energy as well as reducing traffic in the network (by up to 35%) even when the renewable infrastructure differ greatly (each geographical location has its own optimal combination of wind turbine or solar panel infrastructure depending on availability of sunlight and wind).
- Increased use of renewable energy in the network is equivalent to reduction of brown energy; our solution was able to save brown energy usage between 10-55%.
- We identify trade-offs between the amount of renewable energy and caching performance and show how they interact with each other.
- Content traffic to data centers can be reduced by 24-53% using the adequate caching strategy which directly translates into energy savings at the data center.

The paper is organized as follows: Section 2 presents the related work. Section 3 clearly defines the objectives of our solution, and this is followed by Section 4 which describes our proposed approach. Section 5 describes the results of our experiments. Section 6 summarizes the paper, and finally, the appendix describes the meteorological background used for our study, and the mechanisms of converting renewable energy to consumable power.

2 RELATED WORK

In this section, we will present related work in green Internet and information-centric caching, as these are most closely related to our proposed solution.

2.1 Greening the Internet

Developing ideas for a greener Internet has been investigated for a number of years [?], [?], where proposed solutions include new routing approaches as well as considering renewable energy as a possible source of power.

2.1.1 Energy-aware protocols

Bolla et al. [?] developed a new approach for next-generation backbone network devices that have smart stand-by primitives through virtualization of the physical network infrastructure. Their evaluation showed the ability to manage hardware wakeup and standby events transparently from the network-layer protocol. Cianfrani et al. [?] investigated a simple modification of a link-state routing protocol to minimize the number of links powered-on in a topology. This led to traffic being redirected to a small number of nodes, yielding reductions in energy consumption. Solutions based on energy optimization techniques have also been proposed. In [?], an optimization approach was proposed to minimize the traffic and energy consumption of the network. However, these problems are NP-complete which makes them unrealistic solutions for large-scale network. An important requirement is to have a more adaptive [?] power-aware routing mechanism that constantly follows the traffic behavior, while reducing the energy consumption without disrupting the QoS requirements of the network [?].

2.1.2 Renewable energy

Liu et al. [?] proposed the use of renewable energy for powering data-centers, including an optimal mix of renewable sources of energy, using a 30 kW wind turbine and a 4 kW solar panel. The authors found the optimal energy proportion to be 80% wind and 20% solar, which is mainly due to the extra power than can be generated by a 30 kW wind turbine compared to a 4 kW solar panel. Unlike [?], we do not focus on a single data center; instead our goal is to maximize the use of renewable energy to power the routers within the network.

2.2 Content Caching Strategies

Information-centric networking (ICN) [?], [?], [?], [?] has emerged as a general, network-wide caching solution. In ICN the contents are cached in the network (e.g., in routers), where they are able to serve requests that pass through the routers.

A number of research works have investigated and analyzed the performance of ICNs. In [?], Carofiglio et al. provided an analytical model for data transfers in ICN. Muscariello et al. [?] analyzed the ICN performance by taking into account different bandwidth and storage limits. The solution proposed by [?] evaluated the ICN performance with realistic traffic mix, showing there is significant difference between different traffic types

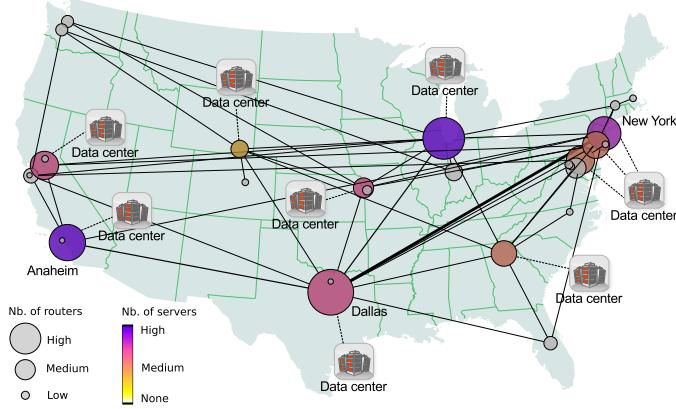


Fig. 1. Sprintlink USA mainland network. The figure shows the network links and the size of the bubbles indicates the size of the POP. The placement of data centers is explained in Section 5.

(Web, file sharing, UGC, VoD, etc.). In all these solutions, only simple Least Recently Used (LRU) approach was used as a caching strategy. Age and popularity-based methods have also been investigated, where solutions have been developed to placing contents closer to the edge caches [?], [?]. In [?], an investigation was conducted on the allocation of cache sizes across the ICN, where results from the analysis showed that the node's cache size should be proportional to its degree. In our previous work [?], we investigated effects of cache admission policies and cooperative caching and found them to be crucial towards good caching performance. The variant of ICN that we consider in this paper is based on [?], which bears a close resemblance to CCN/NDN [?], in that any router on the path of a request may answer it, if it has the content cached.

Our work in this paper differs markedly from the work above. Our focus is not on optimizing caching performance, but instead we combine caching with renewable energy-aware routing and evaluate the combined solution in terms of its ability to exploit renewable energy available at different routers in the network, as well as having a positive influence on the data centers' energy savings.

3 PROBLEM STATEMENT

We now define our problem statement, and models that we use to represent the renewable energy consumption by the routers of the network. We consider an Internet Service Provider (ISP) network (such as Fig. 1), in which nodes may be powered partially from renewable energy sources, where this source could be a combination of wind and solar. The ISP network connects the data centers and the users access points, which serve as the source and destination points for traffic in the network. The remaining routers in the ISP network serve as transport nodes that form the ISP's mesh topology.

An informal description of the design problem we consider is the following:

[Given] (i) a realistic ISP network topology composed of routers and bi-directional links, (ii) an infrastructure for renewable energy for each router, (iii) the power consumption of all devices in the networks besides data centers, and finally (iv) a model for content popularity in function of time.

[Objective] is to (i) determine the impact of integrating renewable energy awareness on the network and the data center performance, and (ii) if needed, determine the best trade-off to maximize the reduction of brown energy for the ISP networks and the data centers, while maintaining satisfactory content delivery performance for the end users.

The energy consumption model of the ISP's network devices is derived from the study of Chabarek et al. [?], which was extended from the model proposed initially by Gupta and Singh [?]. The model proposed by [?] was developed by empirically monitoring the energy consumption of chassis and line-cards in two Cisco routers (*Cisco GSR 12008 and 7507*) under various traffic loads. Results from the experiments showed that the chassis is the biggest consumer of energy while the line-cards consume less.

In this paper, a new energy consumption model for routers has been developed to incorporate use of renewable energy. [?]. The general energy model for a router's power consumption is represented as:

$$PC(X) = CC(X_0) + \sum_{i=1}^N (TP(X_i) + LCC(X_i)) \quad (1)$$

where X is the vector of chassis and line-cards energy models, as well as the traffic configuration of the router. $CC(X_0)$ is the energy consumption for the chassis, N is the number of line-cards of router X , $TP(X_i)$ is the energy consumption due to traffic on the line-card i , and $LCC(X_i)$ is the line-card energy consumption.

Since the impact of traffic load on the routers does not fluctuate significantly, the $TP(X_i)$ can be omitted in equation 1, leading to:

$$PC(X) = CC(X_0) + \sum_{i=1}^N LCC(X_i) \quad (2)$$

In our energy consumption model, we assume that a router has the ability to power-off line-cards, and the entire router will be powered-off when all its line-cards are off. Therefore, the power consumption model at time t can be represented as follows:

$$PC(X, t) = x_{0,t}CC(X_0) + \sum_{i=1}^N x_{i,t}LCC(X_i) \quad (3)$$

where

$$x_{0,t} = \begin{cases} 0 & \text{if } \sum_{i=1}^N (x_{i,t}) = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$x_{i,t}, i > 0 = \begin{cases} 0 & \text{if } X_i \text{ is powered-off at time } t \\ 1 & \text{if } X_i \text{ is powered-on at time } t \end{cases}$$

Consequently, the configuration of each $x_{i,t}$ is added to the vector X . Let X^0 be the vector of chassis and line-cards models, where all $x_{i,t} = x_{i,t}^0 = 1$.

Let us assume that each router X would have a source of renewable energy $rePC(X, t)$ at time t :

$$rePC(X, t) = P_w(X, t) + P_s(X, t) \quad (4)$$

where $P_w(X, t)$ and $P_s(X, t)$ are respectively the power generated from wind and solar energy at time t . More details can be found in the appendix on the methods used to calculate $rePC(X, t)$ using real meteorological data.

By subtracting $rePC(X, t)$ from $PC(X)$, we obtain $brPC(X, t)$, the brown energy consumed by router X at time t .

$$brPC(X, t) = PC(X^0, t) - rePC(X, t)$$

with $brPC(X, t) = 0$, if $rePC(X, t) \geq PC(X^0, t)$ (5)

Our aim in this paper is to analyse the impacts of favoring routers powered by renewable energy in order to facilitate the powering-off of the unused routers that may be powered by brown energy. Therefore, we can formulate the brown energy reduction of the entire network σ_n at time t by calculating the total brown energy consumption of the network when unused devices are powered-on or off.

$$\sigma_n = 1 - \frac{\sum_{\forall X} \sum_{\forall t} brPC(X, t)}{\sum_{\forall X} \sum_{\forall t} brPC(X^0, t)} \quad (6)$$

Let $\alpha, \alpha \in [0, 1]$ be the factor that favor renewable energy, and where $\sigma_n(\alpha)$ and $\sigma_d(\alpha)$ are respectively the brown energy reduction utility functions of the network and the data centers derived from the choice of α . In the subsequent sections we will show how α is used in our renewable energy aware gradient-based routing algorithm. While our paper does not focus on the utility function of the data centers, we discuss the trade-off that can be obtained by changing α and show that the α value can be optimized according to seasonal changes.

To conclude, we want to maximize $\sigma_n(\alpha) + \sigma_d(\alpha)$, while minimizing the impact on the performance for the end users.

4 PROPOSED SOLUTION

Fig. 2 shows an example of our proposed approach, where routes are discovered via routers with access

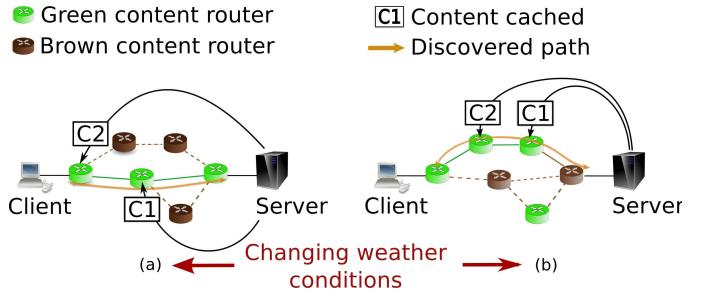


Fig. 2. Illustration of proposed approach, demonstrating path discovery as weather condition changes ((a) to (b)). Once paths are discovered, contents C1 and C2 are cached on routers along the discovered paths.

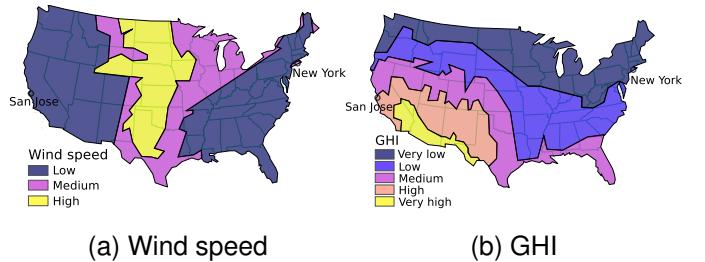


Fig. 3. Average annual wind speed and Global Horizontal Irradiance in USA. Brighter colors indicate higher availability.

to high renewable energy. Once the paths are discovered, contents will be pulled from the data centre and populated along the discovered paths. However, these paths will change as the weather pattern changes, which implicitly moves the contents to new locations that have access to high renewable energy. At the same time, certain routers that are powered by “brown” energy maybe unused, which could lead to a situation where we can power-off these routers to minimize overall brown energy. Therefore, as shown in Fig. 2, the proposed approach is dual-layered.

Firstly, a gradient-based routing protocol discovers the path of routers powered by the highest amount of renewable energy, and secondly a CCN-like [?] caching approach populates the routers of the discovered path with contents based on a caching strategy. Therefore, the approach maximizes the dynamic properties of caching, where contents can be moved to different locations, and in this particular case to locations that provide the highest amount of green energy.

An important factor that affects the performance of renewable energy production is the location. For instance, Fig. 3 shows the average annual wind speed and Global Horizontal Irradiance in the USA. As shown by these two figures, the available energy sources greatly differ depending on the locations of routers. Cities such as New York or San Jose exhibits dissimilar weather conditions, including predictability in the weather patterns.

More detailed information can be found in the appendix which discusses the quantity of wind and solar energy for specific cities. While the production of renewable energy fluctuates for different locations, this is most ideal for our gradient-based routing protocol. The gradient-based routing protocol, which is highly scalable and distributed, can discover routes by adapting to these weather conditions in order to maximize the renewable energy.

The metric used to determine the greenest router is based on the highest ratio of renewable energy consumed compared to total energy requirements of the router $g(X, t)$ (this is also termed *green ratio*). This ratio is calculated according to equation 7:

$$g(X, t) = \begin{cases} 1 & \text{if } rePC(X, t) \geq PC(X, t) \\ \frac{rePC(X, t)}{PC(X, t)} & \text{otherwise} \end{cases} \quad (7)$$

Examples of the variations of the green ratio for different locations are presented in Fig. 14i-14l. Every router in the network locally broadcasts messages to its one-hop neighbors regarding the ratio of renewable energy that it has available. The greenest routers, will in turn maximize the production of packets processed with renewable energy (*green packets*).

The subsequent sections, we describe our routing protocol and caching strategies in more detail.

4.1 Routing

Our routing protocol, called *Renewable energy-aware Parameterized Gradient Based Routing (rePGBR)*, is extended from our original *Parameterized Gradient Based Routing (PGBR)* protocol [?]. The original PGBR is a fully distributed bio-inspired routing protocol, that is inspired by the *Bacterial Chemotaxis* process. Through the process of *Chemotaxis*, bacteria are able to mobilize themselves towards a destination point by sniffing a chemical gradient emitted from the destination node. This same bacteria motility principle is used for the *rePGBR* routing algorithm, where the routes are discovered by hopping from node to node along the path with the highest gradient, until it reaches the destination. The benefit of the *rePGBR* algorithm, includes (i) high scalability, (ii) no requirements for pre-knowledge of the traffic demand, and (iii) the ability to efficiently discover paths to contents favoring the usage of renewable energy.

The gradient field equation is represented as follows:

$$G_i^d(j) = \alpha g(j) + (1 - \alpha)h_i^d(j), \quad 0 \leq \alpha \leq 1 \quad (8)$$

where $G_i^d(j)$ represents the gradient value of the link $i \rightarrow j$ for a packet to destination d , $g(j)$ (e.g. equation 7) represents the green ratio value of neighbor node j , and $h_i^d(j)$ represents the normalized hop count value of neighbor node j of node i for destination d . The α , as specified in Section 3, represents the weighting parameter between the shortest path and the greenest path to the content location, and is a key parameter in evaluating the trade-off between maximizing green energy and

Algorithm 1 Selection of next hop during discovery

Require: Destination d , incoming neighbour iN

- 1: $u \leftarrow$ current node
- 2: $qIn \leftarrow$ list of incoming neighbours
- 3: $qOut \leftarrow$ list of outgoing neighbours
- 4: **if** iN is not nil **then**
- 5: $qIn \leftarrow qIn + iN$
- 6: **end if**
- 7: $nextHop \leftarrow nil$
- 8: $bestG \leftarrow -1$
- 9: **for** neighbor v of u **do**
- 10: **if** $v \notin qIn \cup qOut$ and $G_u^d(v) > bestG$ **then**
- 11: $bestG \leftarrow G_u^d(v)$
- 12: $nextHop \leftarrow v$
- 13: **end if**
- 14: **end for**
- 15: **if** $nextHop$ is nil **then**
- 16: **repeat**
- 17: $nextHop \leftarrow last(qIn)$
- 18: **until** $nextHop \notin qOut$
- 19: **end if**
- 20: **if** $nextHop$ is not nil **then**
- 21: $qOut \leftarrow qOut + nextHop$
- 22: **end if**
- 23: **return** $nextHop$

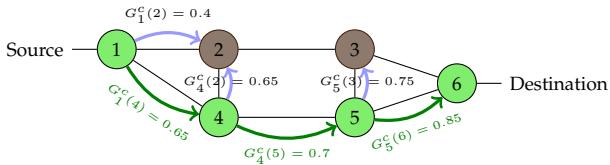
maximizing caching performance. Unlike the normalized hop count value used in [?], where the equation for $h_i^d(j)$ is $h_i^d(j) = 1 - \frac{w_i^d(j)}{W^d}$, with $W^d = \max(w_i^d(k))$, $\forall k$, the new normalized hop count value of *rePGBR* provides fast destination discovery in a scalable manner [?]. The modified $h_i^d(j)$ is represented as:

$$h_i^d(j) = \frac{\max(w_i^d(k)) - w_i^d(j)}{\max(w_i^d(k)) - \min(w_i^d(k))}, \quad \forall k \text{ neighbours of } i \quad (9)$$

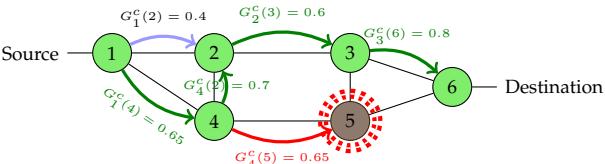
where $w_i^d(k)$ represents the weight of node k in the shortest path tree for root d , the destination.

Routing is performed by first discovering the path between a source-destination pair, by sending a discovery packet that migrates hop-by-hop, selecting the link with the highest gradient value. Algorithm 1 describes how to choose the next hop during discovery. Unlike the original PGBR where the path was stored in the discovery packet, *rePGBR* stores local information in routers. This local information enables the discovery to avoid loops (i.e. using list of incoming and outgoing neighbours, line 10) and performs back-tracking (i.e. lines 15 to 19) in a fully and efficient distributed manner.

Furthermore, this discovery process is depicted in Fig. 4a and Fig. 4b illustrates the discovery process dynamically changing to avoid brown hot-spot areas (e.g. discovery avoids brown node 5 due to the gradient field changes). Once the weather conditions change, a new discovery is issued. However, during the discovery process, the existing path remains stable until the new discovered path is adopted sequentially as the discovery



(a) Discovery of paths along gradient field formed by routers with high renewable energy.



(b) Gradient field modification due to changes in renewable energy performance.

Fig. 4. Route discovery process for *rePGBR*.

message backtracks to the source using the local information described in Algorithm 1.

The utilization of the weighting parameter α provides *rePGBR* with robustness in supporting multiple objectives. For instance, when α is set to a low value (e.g. 0.2), the discovered path is biased towards the shortest routes (more weight is attributed to the hop count) while a larger value would tend to favor greener routes. As a result, the gradient field equation of *rePGBR* is flexible and can exhibit multiple behaviors by only manipulating one parameter to provide the desired green performance of the network. This paper aims to investigate the most optimum α , that could maximize the savings of brown energy usage of the full network, as well as having a positive influence on the data centers' energy savings.

4.2 Caching

This section will describe the second component of our solution, which is the caching process of contents along the paths discovered by *rePGBR*. We used three caching strategies described in [?]. The work in [?] focused on evaluating these strategies purely on caching performance across a wide range of networks. In this paper, we focus on using these caching strategies to maximize the use of renewable energy for specific points of storage, and the suitability of these strategies in combination with the adapted *rePGBR* routing algorithm. Below we describe the three strategies and briefly recap the main results from [?]. In terms of different information-centric architectures, our caching strategies would be most suitable for a CCN/NDN-like architecture [?].

ALL: ALL is the simplest caching strategy, where each router tries to cache all the packets that passes through, and uses LRU algorithm to evict packets if the cache is full. There is no cooperation between the routers.

Cachedbit: Cachedbit uses one bit in the packet header to indicate whether the packet has already been cached by a router along the path. If the bit is not set, a router

Algorithm 2 Admission Policy in Cachedbit Strategy

```

1: Input: Data chunk  $C_i$ 
2: Output: Caching decision
3: if  $C_i$  is not cached and cached_bit not set then
4:   Draw a random  $x$  uniformly from  $(0, 1)$ 
5:   if  $x < 1/n$  then
6:     if Cache is full then
7:       Evict entry based on LRU
8:     end if
9:     Add data chunk  $C_i$ 
10:    Set cached_bit
11:   end if
12: end if
13: Forward chunk to destination

```

Algorithm 3 Cooperation Policy in NbSC Strategy

```

1: Input: Data request  $R_i$ 
2: Output: Response decision
3: if  $C_i$  is cached for  $R_i$  then
4:   Reply with  $C_i$ 
5: end if
6: if  $C_i$  is not cached for  $R_i$  then
7:   if Neighbor  $N_i$  cached  $C_i$  for  $R_i$  then
8:     Redirect  $R_i$  to  $N_i$ 
9:   else
10:    Forward  $R_i$  to next hop
11:   end if
12: end if

```

decides locally whether to cache the packet or not, with a probability of $1/n$ where n is the length of the path from the client to the server (or ingress router to egress router for an intermediate network). As discussed in [?], the length of the path can be determined by observing packets that pass through a router; global knowledge is not required. In this case, the caches use LRU as a replacement policy for evicting the contents. Algorithm 2 describes the admission policy for the Cachedbit strategy.

Neighbor Search (NbSC): NbSC is otherwise similar to Cachedbit, except that routers periodically send Bloom filters of their contents to their neighbor routers. When a router experiences a miss, it can check if any of its neighbors has the requested packet. If there are multiple matches, the request will be redirected to a random neighbor. We also investigated a variant which redirects requests to the greenest neighbor.

The results in [?] showed that a Cachedbit-like admission policy is needed to get good caching performance, but that the addition of NbSC reduces network traffic considerably.

Algorithm 3 describes the cooperation policy in the algorithm that implements NbSC.

5 EXPERIMENTATION

In order to validate our proposed approach, we have conducted experiments on a testbed using the weather data described in the appendix. The network evaluated on our testbed is based on the Sprint router-level topology from the Rocketfuel project [?]. The Sprint network consists of 278 routers geographically distributed in 27 cities in U.S. mainland (meteorological data from [?] is available only for the U.S. mainland locations). For the experiment, the top 40 highest degree nodes are connected to content servers, while the 80 routers with the lowest degrees are connected to the clients. Fig. 1 shows the resulting topology where routers located in the same city are grouped (i.e. the size of the city increases with the number of routers). For instance, Anaheim, Chicago, New York and Dallas are the cities with the highest number of routers. The color of the cities, on the other hand, represents the number of servers (e.g. nodes which have the highest connectivity). A total of 10 cities have servers resulting in 10 data centers that are interconnected through the ISP network. Once again, Anaheim, Chicago and New York are the three cities with the highest number of servers.

The remaining routers are used as intermediate nodes for the routing algorithm. In our experimental setup, the clients will continually request data from the servers (request rate depends on the traffic pattern). While tests have been conducted for three different traffic patterns, which includes (i) real trace from the TOTEM project [?], (ii) trace from the gravity model [?], and (iii) constant traffic pattern, there was no significant difference between these traffic patterns in terms of performance. Therefore, we only show the results of our experiments based on a constant traffic pattern.

The request trace for each client is generated from a real DNS trace of a university lab. We sorted the requested DNS names according to their popularity and removed the top 20 entries. The request pattern from this trace follows a Zipf distribution with parameter 0.9, which is very close to real-life distribution shown in [?]. Our trace requests chunks of content, which are assumed to be independent of each other. We assigned each router with a storage capacity of 4 gigabytes. The power requirements of every router has been set using the measurements of a Cisco 7507 [?]: the basic chassis requiring 210W and each line-card an additional 70W (i.e. a router with 4 line-cards consumes $210 + 4 \times 70 = 490$ W).

We investigated two scenarios: In scenario A, the renewable energy infrastructure of the routers are set to supply twice the energy required by the router (i.e. $rePC(X, t) = 2PC(X^0, t)$) at most. Therefore, the changes of weather conditions greatly modify the green ratio $g(X, t)$. This scenario is defined as the static capacities scenario. In scenario B, the infrastructure supplies up to three times the energy required by the routers. The different size of infrastructures implies that some routers are mostly fully brown and others mostly green,

which will better represent a realistic scenario where green renewable energy farms are built gradually (to limit the cost of installation). More details on the settings can be found in the appendix.

The following metrics have been used to evaluate our proposed solution for the two scenarios previously described:

- **Hit Rate:** Determines the fraction of requests that are served by the content routers. Since all objects are of the same size, the hit rate is also the byte hit rate. Hit rate shows how much external traffic is saved by caching, thus leading to potential energy savings for the data centers.
- **Footprint Reduction:** Network footprint is the product of the amount of content and the network distance from which the content was retrieved. It measures the amount of internal traffic reduction, where a smaller footprint (larger reduction) means less traffic within the network.
- **Green/brown ratio:** Proportion of the packets by the router within an hour that has been processed using renewable energy to the packets that has been processed using fossil fuel energy.
- **Reduction of brown packets:** Reduction of brown packets that has been processed within an hour using fossil fuel energy.
- **Brown energy savings:** The amount of brown energy that has been saved by powering-off unused routers. In the case of a neighbouring router being unused, the line-card connecting the two routers is considered to be powered-off. The brown energy savings is calculated using equation 6.

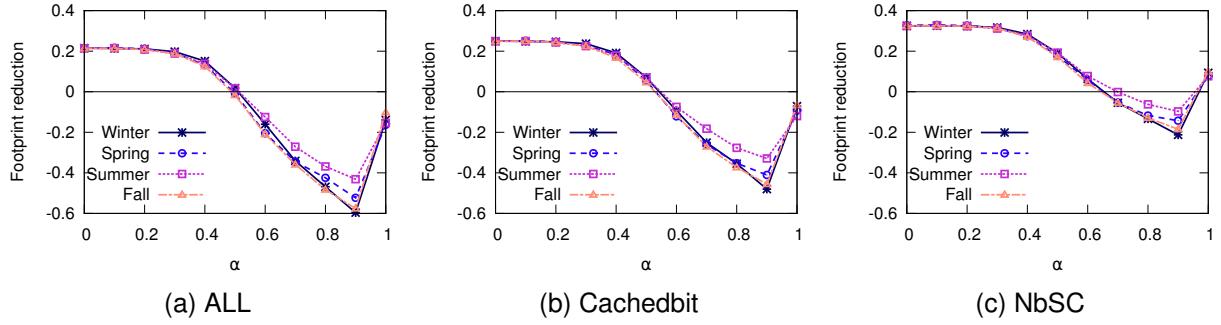
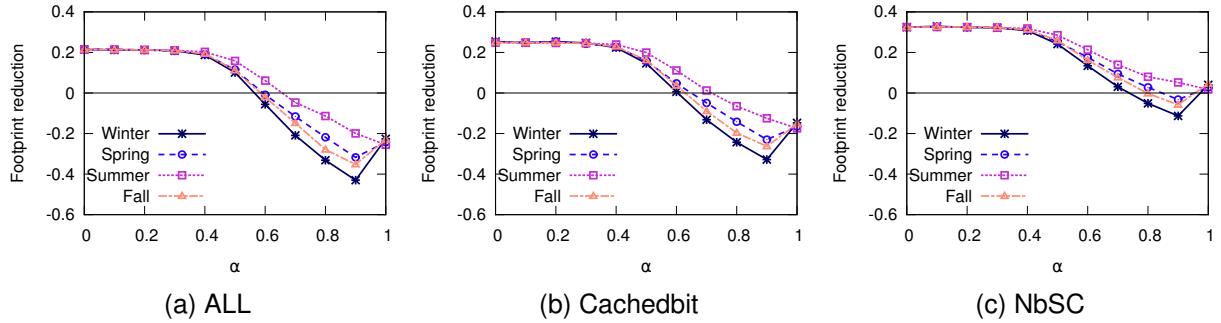
While hit rate and the footprint reduction are mainly caching metrics, the green/brown ratio, reduction of brown processed packets and the brown energy savings are representing the performances of the solution to successfully address the energy efficiency requirements. Hit rate also represents the reduction on traffic towards data centers and, therefore, indicates potential energy savings at the data center.

5.1 Testbed setup

All the experiments are performed on a cluster of 240 Dell PowerEdge M610 nodes. Each node has 2 quad-core CPUs, 32GB memory, and is connected to a 10-Gbit network. All the nodes run Ubuntu SMP with 2.6.32 kernel. The experimental platform we used in the evaluation is capable of simulating realistic routers, and allocating necessary physical resources according to the simulated network size. In the event that the network size is larger than the actual number of nodes in the cluster, multiple routers will be multiplexed onto one node.

5.2 Network performance

As described earlier, the benefit of the *rePGBR* routing algorithm is in the ability of manipulating α to suit

Fig. 5. Varying α of *rePGBR* routing and evaluating its impact on footprint reduction for scenario A.Fig. 6. Varying α of *rePGBR* routing and evaluating its impact on footprint reduction for scenario B.

any particular objective. In our scenarios, we would ideally like to select an α that provides the greenest routes, without compromising on the caching performance. Therefore, we have conducted tests to select the most appropriated α for different seasons in the year. In total, we investigated 4 weeks throughout the year, and the periods of these weeks are as follows (Winter: 1st - 7th, January; Spring: 6th - 12th, April; Summer: 11th - 27th, July; Fall: 21st - 27th, October). The weeks of Spring and Fall have been selected at different times to ensure that the weather patterns would not be similar. Fall is on the verge of Winter while Spring is slightly milder.

We first evaluate the different caching strategies with respect to varying α values for scenarios A and B. Below we will look more closely at performance of *rePGBR* in these conditions.

We begin first by evaluating the different caching strategies with respect to varying α values. As previously mentioned, we focused on 5 metrics including footprint reduction (e.g. Fig. 5 and 6), hit rate (e.g. Fig. 11), green/brown ratio and the reduction of brown processed packets (e.g. Fig. 7), as well as the brown energy savings of the network (e.g. Fig. 8 and 9). All the results are shown using 5 kW wind turbines, but with the nodes having different renewable energy architectures (scenarios A or B). Below we will look more closely at the performance of the solution in these conditions.

5.2.1 Footprint reduction

We previously mentioned that the footprint reduction is a relevant metric for the ISP, showing the improvements in its internal traffic. Fig. 5 shows the footprint reduction as a function of α for the four seasons and for each caching strategies. For all the caching strategies, we see that as α increases, the footprint reduction decreases. First the degradation is limited (when $\alpha \leq 0.3$). The reason for this positive correlation is because higher α value leads to longer paths along routers with high green ratio, which further increases the footprint of the network. The results also show NbSC is far more superior than Cachedbit and ALL. The same behavior is observed for scenario B and Fig. 6 confirms this.

In particular, after α reaches 0.4, the footprint reduction drops significantly. When the footprint reduction reaches zero, this means the paths are so long that caching strategies can no longer reduce intra-network traffic. Obviously, caching still happens, but the net benefit is zero. From the figure, we can see this threshold is about 0.8 for NbSC (e.g. Fig. 6c) and lower for the others (e.g. Fig. 6a and 6b). The degradation of the footprint reduction varies with the seasons but the behavior is similar. Although the behaviour is similar, we can see that for all caching strategies there are improvements in footprint reduction in the warm seasons. This improvement is visible when α is greater than 0.6 and 0.4 for scenario A and B, respectively. An increase of the footprint reduction happens when $\alpha = 1$, due to inefficient route discoveries that are not able to update routing table

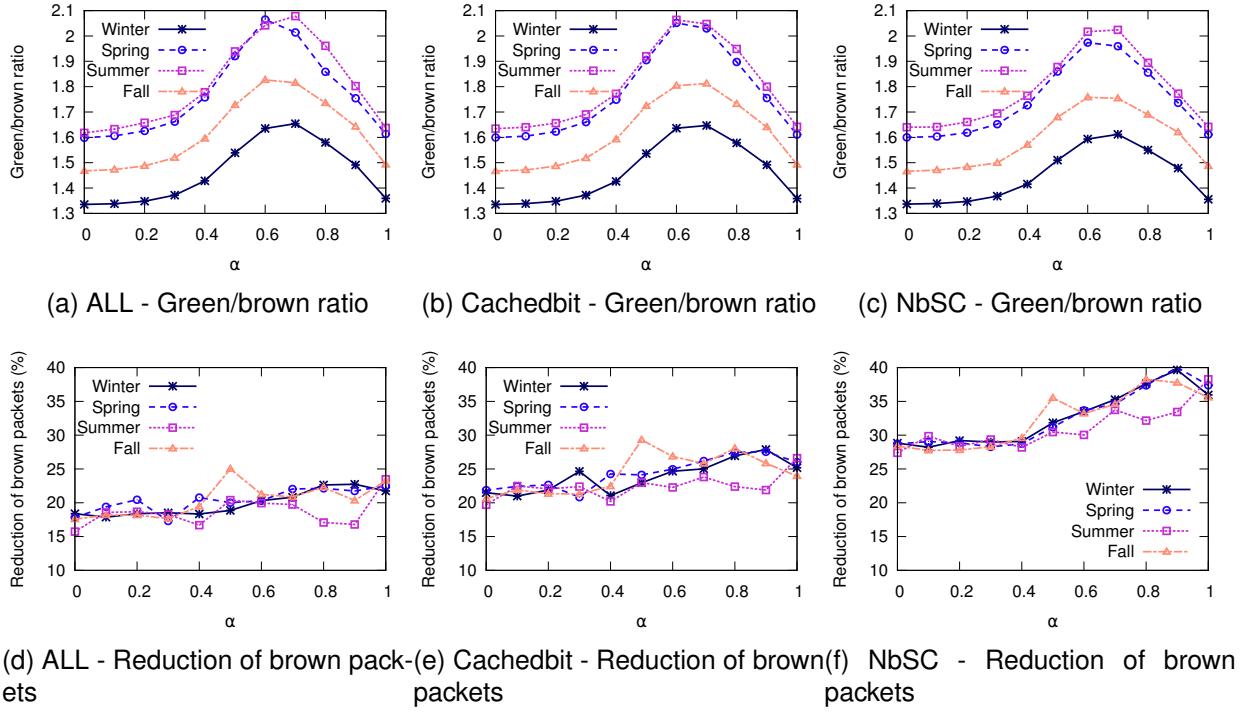


Fig. 7. Varying α of *rePGBR* routing and evaluating its impact on green/brown ratio and reduction of brown packets for scenario A.

information with the latest renewable energy values. As a consequence, requests are served using outdated paths increasing slightly the chances of a cache hit.

As we mentioned before, we considered a variant of NbSC which selects the neighbor router with the highest green ratio, instead of a random neighbor when multiple matches are found. However, we found out that this modification has no visible effect on the performance. This is mainly because routers are co-located at POPs and thus share the same weather pattern among many (but not all) neighbors.

We observe a large performance difference between NbSC and the others even when using only the shortest path (with $\alpha = 0$; ALL and Cachedbit has a footprint reduction of 21% and 25%, respectively, while NbSC has a footprint reduction of 32% for the static and random scenarios). The advantage of NbSC is at least 7% higher than Cachedbit and increases with α , up to 23% and 18%, respectively for scenarios A and B.

5.2.2 Improving network's greenness

Fig. 7 shows the ratio of green/brown packets and the reduction of brown packets as a function of α for scenario A. In particular, Fig. 7a - 7c show a difference between the seasons for the green/brown ratio but very little between the caching strategies. This is due to the importance of *rePGBR* route discovery to enhance the green processing of the packets. In addition, we can see that the ratio is greater for the warm seasons (i.e. Summer and Spring) than for the cold seasons. One

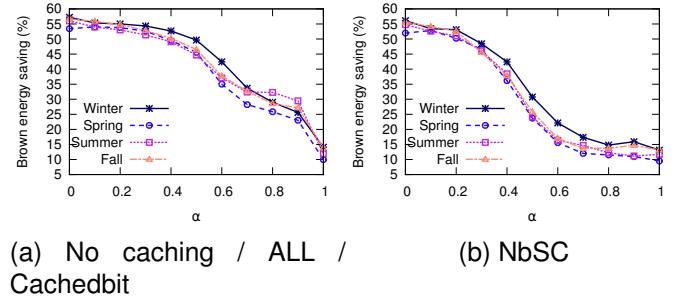


Fig. 8. Varying α of *rePGBR* routing and evaluating its impact on brown energy savings for scenario A.

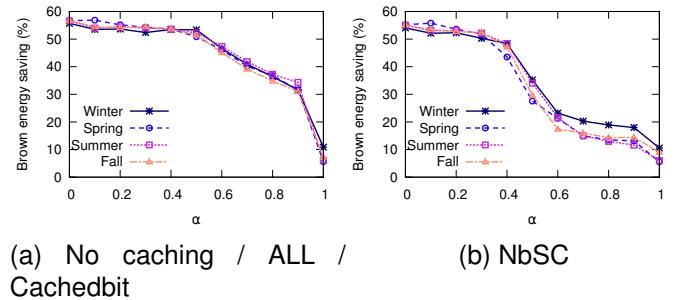


Fig. 9. Varying α of *rePGBR* routing and evaluating its impact on brown energy savings for scenario B.

can see that the performance of the warm seasons are quite similar compared to those for the cold seasons.

As explained in the appendix, the warm seasons tend to be more reliable (the impact of the solar energy is higher than the wind energy) and predictable (diurnal phases) which leads the routing protocol to maximize the discovery of highly green routes. Fig. 7a - 7c show that the ratio of green/brown packets slowly increases with α . As shown in the figures, the significant improvement (peak) starts at approximately 0.4 and ends at 0.9. Therefore, this is the range of α values that maximize the green processing of packets in the network. However, the ratio of green/brown packets is insufficient without showing the reduction of brown packets as a function of α , which are shown in Fig. 7d - 7f. The results show that the number of brown packets almost linearly decreases as α increases, which comforts the fact that the network's greenness increases when $0.4 \leq \alpha \leq 0.9$. Finally, the results also show that NbSC outperforms other caching strategies for every α .

Experimental results show similar trends for scenario B, that differ only from the green/brown packet ratio induced by the scenario definition. In the appendix, we described how the capacities influence on the greenness of the routers, where certain routers are more green than others irrespective of the weather conditions. This in turn limits the variety of routes discovered by *rePGBR*.

Using the ability of *rePGBR* to maximize the discovery of green routes, could leave to a number of routers left unused. Fig. 8 and 9 shows the brown energy savings that could be achieved if these routers were powered-off. The best savings are obtained when *rePGBR* goes for the shortest routes, and decreases as α increases. The performances of ALL, Cachedbit and no caching are similar because all the routers discovered by decreasing α are used independently of the caching strategy. However, NbSC presents a significant disadvantage where the brown energy savings are reducing much faster with regards to α . For Cachedbit, the brown energy savings are quite stable until α reaches 0.4, while NbSC maintains high energy savings only until 0.2. To summarize, the brown energy savings of the network varies between 10% and 55% which could lead to significant improvement of the network's greenness. For scenario B, the values of α that maintain high energy savings are larger than for scenario A. For instance, Cachedbit is able to maintain brown energy savings over 50% when α is lower or equal to 0.5 and NbSC to 0.3 for all seasons. As shown in Fig. 8 and 9, the impact of seasonal weather conditions is not significant on the brown energy savings, because these savings are highly dependent on *rePGBR* route discoveries.

5.2.3 Hourly performance

Fig. 10 shows the 4 metrics described above for Cachedbit and NbSC as a function of time for the first 48 hours of Spring for scenario B. The α values used for each caching strategy correspond to the greenest α maintaining brown energy savings above 50%, which are respectively 0.5 and 0.3 for Cachedbit and NbSC (Fig. 9).

In detail, Fig. 10a shows the footprint reduction of Cachedbit and NbSC. The results of ALL behave similarly to Cachedbit's, although yield less gains, and as a consequence have been omitted. The results are showing the high stability of NbSC which maintains a high footprint reduction of 32% while Cachedbit's performance are highly variable throughout the days. The performances of the two caching strategies are similar for the brown energy savings (e.g. Fig. 10b) and the green/brown ratio (e.g. Fig. 10c) and distinctively show a trend that maps to the sun pattern (see appendix). In other words, the performances of the caching strategies drastically increase during the sunny hours. As previously mentioned, the sunny hours bring stability in the route discovery, thus improving the solution's performances. Lastly, Fig. 10b shows the reduction of brown packets as a function of time, where the daytime pattern is not visible, and performances of Cachebit and NbSC are fluctuating respectively around 20.5% and 26.5%.

5.3 Impact of the solution on data centers

In order to improve energy savings in the data centers, we assume that a non-negligible improvement could be obtained by reducing the number of requests directly served by the servers of the data centers. Unlike the routers' energy consumption where the workload induced by traffic does not modify the energy consumed, the data centers energy consumption are known to *vary with their workloads*. Our solution, combining the energy savings at the ISP using greener routes and efficient caching strategies, could greatly improves the workload of the data centers by serving content cached in the ISP's routers, thus reducing their energy requirements (i.e. maximizing $\sigma_d(\alpha)$ defined in Section 3).

We see that as α increases, the hit rate also increases, as shown in Fig. 11. The reason for this positive correlation is because higher α value leads to longer paths along routers with high green ratio, which further increases the possibility of cache-hit along the path. The results also show NbSC is far more superior than the Cachedbit and ALL, with a respective gains of over 15% and 20% improvement for scenario B. Experiments for scenario A have shown similar gains and, thus have been omitted. Furthermore, the seasonal differences are also negligible, which shows that the effectiveness of the caching strategies defined by α would deliver predictable performances.

The hourly performances of Cachedbit and NbSC to improve hit rate are depicted in Fig. 12 for a selection of α values on the two first days of Spring. The α values have been selected to show the hit rate performances when (i) the greenness of the ISP network is at its peak (where α equal to 0.3 and 0.5 for NbSC and Cachedbit, respectively) or when (ii) the hit rate performance is at the maximum (with $\alpha = 0.9$ for both strategies). Fig. 12 shows that hit rate performances are very stable and do

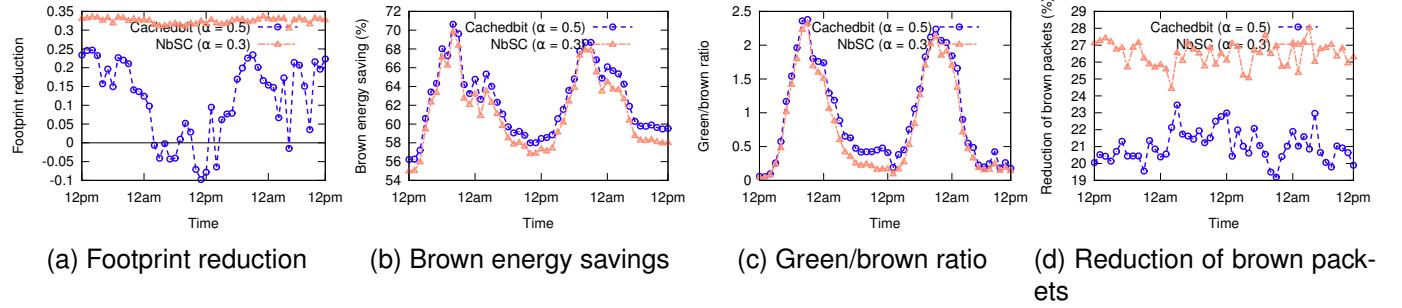


Fig. 10. Performances of Cachedbit and NbSC with respect to time during the two first days of Spring for scenario B.

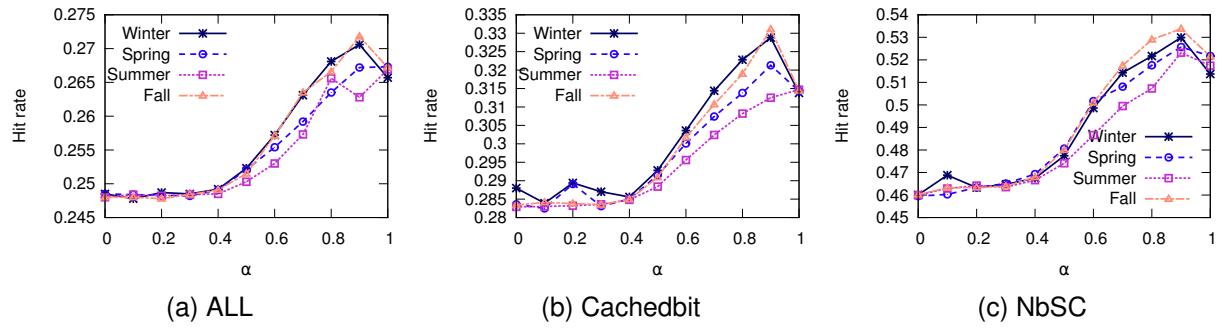
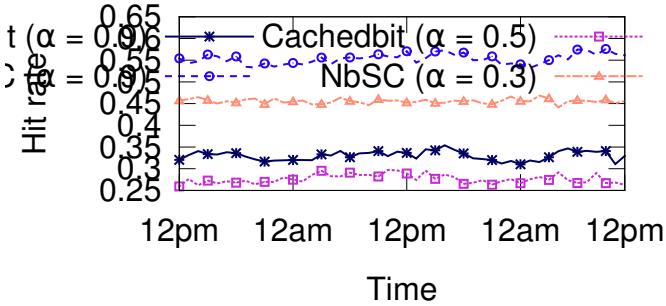
Fig. 11. Varying α of rePGBR routing and evaluating its impact on hit rate for scenario B.

Fig. 12. Hit rate performances of Cachedbit and NbSC with respect to time during the two first days of Spring for scenario B.

not vary according to the weather conditions. However, one can observe that Cachedbit is constantly performing worse than NbSC with a loss of 20% on average in both cases (i and ii).

As a result, increasing renewable energy awareness in routing enhances caching performances of the network, thus resulting on a positive impact on the data centers' workload. This in turn could lead to data centers reducing their energy consumption for unused servers that hold contents which are already cached within the network.

5.4 Discussion

As stated in Section 3, the aim of the paper is to demonstrate that using renewable energy awareness in routing at the ISP's network could reduce the energy consumption from fossil fuels of the network and the data centers serving user's requests. The network operator, could with the choice of a single parameter α and a caching strategy, maximize the brown energy savings of the network from an end-to-end perspective (i.e. including data centers), without compromising the performance of end user's access to contents.

- **User-observed content delivery performance:** In this paper, we assume that delivery performance is positively correlated with footprint reduction. Fig. 5 and 6 showed that the performance would decrease as the renewable energy awareness increases, and this is observable for all caching strategies. There are even α values where the performance is worse than not using any caching (i.e. footprint reduction below 0). For all caching strategies, the delivery performance is maximal when the energy awareness is nil. Fortunately, the performance is maintained for lower α values, and using NbSC further maximizes the delivery performance.
- **Network greenness:** Maximizing network's greenness includes minimizing use of brown energy by powering-off unused routers and reducing the number of brown packets. Unfortunately, these performances are not achievable using the same α values and caching strategies. On the one hand, in order to

maximize the brown energy savings of the network, the smallest α values and Cachedbit are preferable. On the other hand, maximal reduction of brown packets is obtained with high α values and NbSC. Depending on the requirements of the network, the operator would have to carefully choose the appropriate combination to best fulfill the network's objectives.

- **Data center greenness:** Improving caching performances also improves the hit rate and thus helps reduce the workload on data centers. Increasing α improves the hit rate, while using NbSC always increases the performances by an additional 20%.

Therefore, the choice of the caching strategy and α value could produce a positive trade-off, where benefits would be maximized.

6 CONCLUSION

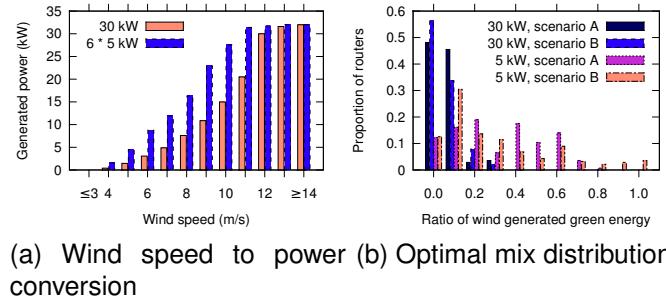
The popularity of the Internet today has led to widespread deployment of ICT infrastructures, which is consuming considerable quantity of energy. In recent years, a new research initiative towards enhancing renewable energy aware ICT infrastructure has been proposed. In this paper, we propose a renewable energy aware CCN, where routers are powered directly by wind and/or solar energy. The proposed approach is composed of two components, which includes a novel gradient-based routing algorithm that discovers paths along routers powered by high renewable energy, followed by a caching strategy that stores contents along these discovered paths. This enables the routing protocol as well as caching strategy, to adapt to varying weather patterns that may affect energy that is used to power the routers. The results from an experimental testbed using real meteorological data, has shown that using the combined approach has resulted in increased consumption of renewable energy, where a combination between high renewable energy consumption and caching strategy can minimize the need of fossil fuels energies of the network and the data centers without compromising the users' experience.

APPENDIX RENEWABLE ENERGY BACKGROUND

We use publicly available data from the National Renewable Energy Laboratory [?] to estimate the availability of wind and solar power in different parts of the U.S. Below we describe our weather data and present the methodology used to convert the meteorological wind and solar data, to power.

Wind energy profile

Fig. 3a shows a simplified version of the annual wind speed in the U.S. The darker zones have an average wind speed between 4 - 6 m/s, the average wind speed for the



(a) Wind speed to power (b) Optimal mix distribution conversion

Fig. 13. Wind energy profile for different turbine types (5 kW and 30 kW), as well as distribution of routers with optimal ratio of wind and solar energy sources.

medium zones are between 6 - 8 m/s, and in the brighter zone, the average speed is greater than 8 m/s.

We consider two wind turbines produced by the Huaying company [?]. The two models are "HY30-AD11" which is a 30 kW wind turbine, and a smaller model "HY5-AD5.6" which produces energy up to 5.4 kW (the larger turbine has a similar size to the turbine investigated in [?].)

Fig. 13a presents a plot showing the relationship of wind speed to generated power that can be achieved by the two different turbines. As we can see, six 5 kW wind turbines produce more power than a single 30 kW turbine at a low wind speed (e.g. at 6 m/s, the ratio is 2.9). However, the type of turbines could be decided on the space restrictions and the cost of the infrastructure.

Solar energy profile

The Global Horizontal Irradiance (GHI) is used to estimate how much power could be generated by photovoltaic solar panels. A 4 kW photovoltaic solar panel that is needed to produce 4 kW of power, will require a GHI of 1000 W/m². As shown in Fig.3b, the GHI is usually stronger in the south west of the U.S. For example, in the brighter zone (very high GHI), the estimated energy is greater than 6000 kWh, while the estimated energy is only 3000 kWh in the zone with the lowest GHI.

Renewable energy calculation

In this section, we describe two calculations that determine the required renewable energy infrastructure to provide adequate renewable energy (the infrastructure refers to the number of turbines or size of photovoltaic panels). Calculation 1 will be used when the proportion between wind and solar power is predefined. Calculation 2 will determine the optimal proportion of wind and solar power that maximizes the green ratio of the router.

Calculation 1: Given a proportion β , a capacity c and $PC(X^0, t)$ the energy required by the router X , we intend to determine the renewable energy infrastructure,

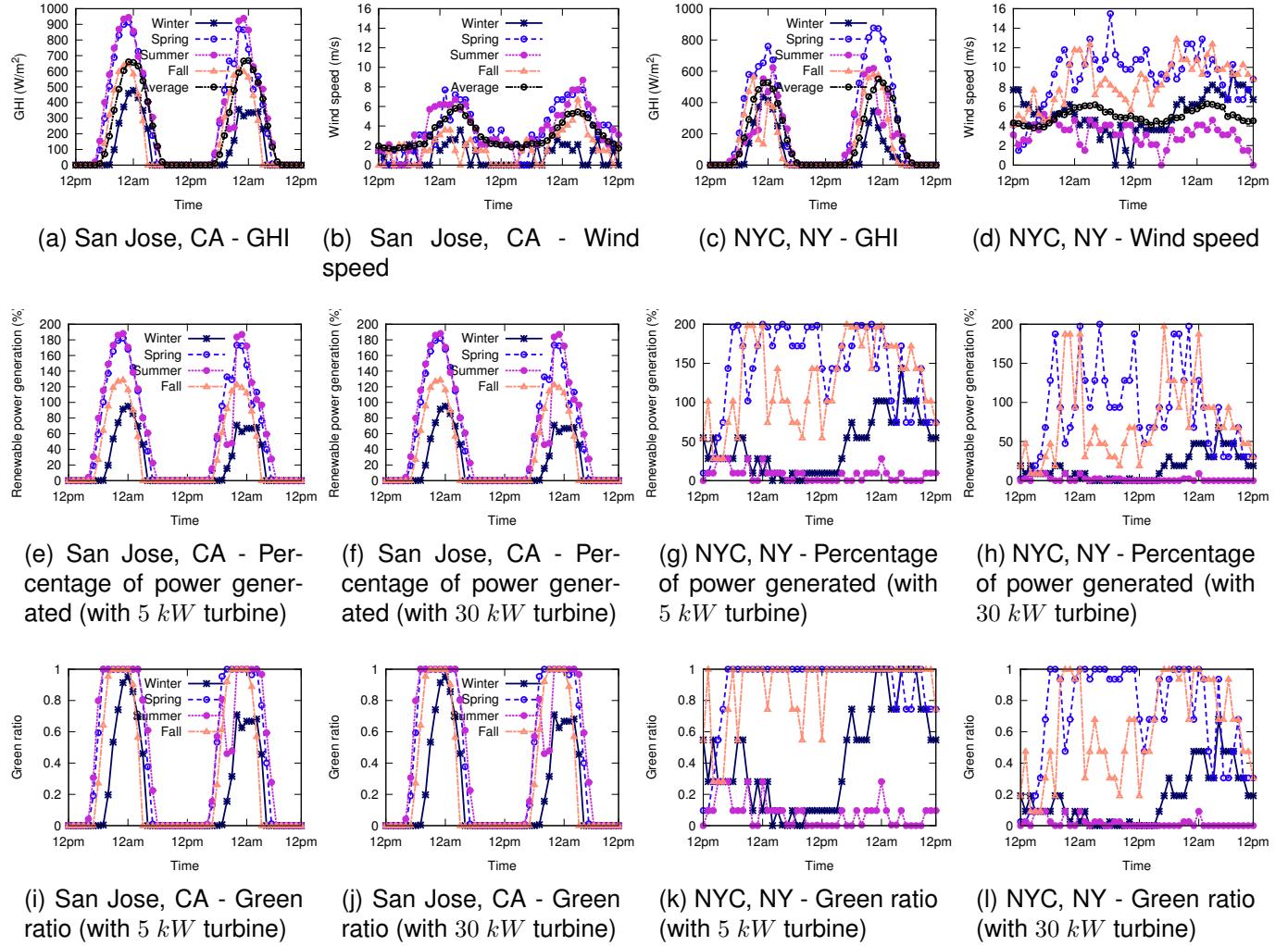


Fig. 14. Weather profiles and power generation performance for San Jose, CA and New York City, NY (The number of turbines has been optimized by calculation 2).

such that:

$$rePC(X, t) = P_w(X, t) + P_s(X, t), \\ \text{where } P_w(X, t) = \beta rePC(X, t) \quad (10)$$

In this case, $P_w(X, t)$ is the energy provided from wind infrastructure, while $P_s(X, t)$ is the energy provided by the solar panels. Therefore, $rePC(X, t)$ is the total renewable energy that can be provided to router X . At the same time we also consider cases when a large quantity of renewable energy is produced. In such situation, the capacity c is set to satisfy $\max(rePC(X, t)), \forall t = cPC(\bar{X}^0, t)$.

Calculation 2: In order to maximize the green ratio of the router, we decided to find the optimal β that would maximize the average $g(X, t)$ ($avg(g(X, t))$), for a typical year. The $avg(g(X, t))$ is calculated as follow:

$$avg(g(X, t)) = \frac{\sum_{t=1}^{24*365} g(X, t)}{24 * 365} \quad (11)$$

where $g(X, t)$ is the green ratio of router X at time t defined by equation 7 (time has been shifted to match

the PDT time zone of Los Angeles, CA). We will do this calculation for all β between 0.0 to 1.0. Therefore, the optimal β will be the value with the best $avg(g(X, t))$.

Fig. 13b shows the distribution of routers for the Sprint network topology as a function of the optimal β . For example, in Fig. 13b, when no wind energy are considered in the ratio (i.e. $\beta = 0.0$), for a maximum $c = 2$ for 30 kW wind turbine, we can see that only 47% of the routers will have this configuration. Fig. 13b shows the optimal mix for both types of turbines for two different types of capacities. The two different capacities, includes a fixed capacity $c = 2$ (also referred to as constant capacity for the scenario A) and random capacity that is uniformly selected between 0 and 3 for the scenario B. As shown in Fig. 13b, selecting the 5 kW turbine will lead to a better distribution of routers with mixed sources of wind and solar (where β goes up to 1.0). On the other hand, the 30 kW turbine will only lead to low ratios of wind energy (e.g. $0.0 \leq \beta \leq 0.3$).

We now show an example of the optimization process for two routers, one in San Jose, CA, and the second in

New York City, NY. Fig. 14a and 14c show the GHI and wind speeds for these three locations over a period of two days, for each season. Fig. 14e, 14f, 14g and 14h show the power generated at these two locations for wind energy using either 5 kW or 30 kW turbines, after the optimization of ratio between wind and solar energy. Recall that the optimization is performed over the whole year by picking out the peak periods of high wind and solar performance, and the figures presented only show performance for two days. Interestingly, the router in San Jose gets all of its renewables from solar and ignores the weak wind performance, even though this means that during night time the router must be powered with brown energy. In contrast, the router in New York uses only wind energy as its renewable source (although not clearly visible in Fig. 3a, New York is in the medium wind zone). Fig. 14g and 14h, shows that using 5 kW wind turbines produces better energy performance than 30 kW wind turbines. As a consequence, only experiments using 5 kW wind turbines have been presented in this paper.