Using SDN for Controlling the Carbon Footprint of The Internet

Md. Mohaimenul Hossain, Jean-Philippe Georges, Eric Rondeau, Thierry Divoux

Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France

(e-mail: firstname.lastname@univ-lorraine.fr)

Abstract: Several works have been done in order to balance the energy consumption of the network with traffic which aims to have a positive impact on the CO₂ emission. However, CO₂ and energy consumption cannot be considered proportionate if the means of electricity production differs. In this paper, we have proposed two different metrics namely Carbon Emission Factor and Non-Renewable Energy usage percentage for achieving green network. We have designed an algorithm considering these metrics as objective functions. We have considered a software defined network approach and provided a set of data and control plane for each metric. Their performances are then analyzed and compared with respect to green policy enabled Shortest Path First algorithm. All the experiments are conducted on GÉANT network with realistic demand size. A comprehensive analysis of the quality of service parameters like the end to end delay and packet loss is also done for each metric of the algorithm.

Keywords: Telematics, Green Networking, Carbon Footprint, Carbon Emission Factor, Renewable Energy, Software Defined Networking, Genetic Algorithm.

1. INTRODUCTION

The ease of availability of inexpensive devices, increasing acceptance of different Internet-based services among the users, Internet of things technologies, all these things result in a rapid growth of the Information and Communication Technologies (ICTs). But this massive advancement comes with a price. ICTs are starting to have a non-negligible effect on global warming. Webb (2008) in smart-2020 report mentioned that 2% of the global greenhouse gas (GHG) emissions annually is caused by ICT, a value that surpasses the GHG emission by the aviation sector. Comparable sort of claim is also detailed in digital agenda report (2013) provided by European Commission, “ICT products and services are currently responsible for 8 to 10% of the EU’s electricity consumption and up to 4% of its carbon emissions”.

Engineers and researchers are working together to reduce this adverse effect of the network devices and infrastructures and have proposed several green networking solutions discussed in related work. The key idea behind green networking is to improve energy efficiency of the network which will eventually reduce the carbon footprint. Energy efficiency technique relies on the fact that network infrastructures are often over provisioned as mentioned in Maleki et al. (2017). That is because these networks are designed to perform sufficiently during peak hours. However, traffic varies vividly during different parts of the day. Given this scenario, Energy Aware Routing (EAR) is an efficient and widely used strategy to reduce the consumption of the network. EAR is executed using different energy optimization technique like turning off unused network devices, put them in sleeping mode or Adaptive Link Rate (ALR). Even though introducing of these energy aware routing algorithms increase consumption in controller (i.e. CPU and memory) but compare to the overall savings this amount is negligible. However, most of the green networking solutions are solely focusing on reducing the energy consumption disregarding the fact that, different means of generating electricity produces different level of carbon footprint. The same amount of energy production can have a very different impact on the environment in terms of carbon emission. For example, according to Brander et. al. (2011), if we compare between France and Poland, for producing 1Kwh of energy Poland emits almost 15 times more CO₂ than France. This clearly indicates a similar network device can have a completely different environmental effect based on its energy source. For a globally distributed network, energy consumption cannot be an all conclusive parameter while designing a green network. Moreover, in last few years, data traffic over network has been drastically increased. According to the report of Cisco VNI (2019), the global Internet traffic of 2022 will be equivalent to 75 times of the traffic of 2007. Therefore, in future, networks might not be as over provisioned as they are currently.

In this work, we have proposed two new green metrics for designing a sustainable network. The first one is Carbon Emission Factor (CEF) and the second one is Non-Renewable Energy usage percentage (NRE). We have considered two green policies: shutting down unused nodes and using of ALR when possible. In case of EAR, while using these policies the goal is to minimize the energy consumption whereas in our case, we have changed the objective function based on our introduced metrics. For CEF the goal is to shut down nodes and links and implement ALR in a way so that it reduces the overall CO₂ emission of the infrastructure. CEF addresses the carbon emission problem more directly. On the other hand, some energy resources have a different impact on environment rather than only carbon emission, such as nuclear power plant produces nuclear wastage. Therefore, our second environmental metric is NRE which will be applied to reduce the non-renewable energy usage percentage while considering the same green policies. Our both parameters directly address the carbon footprint of the network.
We have implemented our solution from the standpoint of software defined network (SDN) which is a well-suited architecture for this context. The centralized routing decision-making mechanism provides a co-ordinated approach for applying green routing which can be easily implemented and maintained. The paper is structured as follows: the next section contains the related work. Section 3 contains the proposed solution with problem formalization, section 4 covers evaluation of the solution and section 5 gives conclusions.

2. RELATED WORK

Recently, energy aware routing with SDN has received significant attention in the research community. Heller et al., (2010) have proposed elastic tree for datacentres where they have formalized the common idea of turning off links and switches based on the amount of traffic load. They wanted to make energy consumption proportional of the network and it will change dynamically according to traffic. Similarly, Bolla et al., (2013) have presented the GreenSDN approach, which integrates three different protocols that operate at different layers of the network: Adaptive Link Rate (ALR), Synchronized Coalescing (SC) and Sustainability-oriented Network Management System. Unlike Heller et al. (2010) they have considered both shutting down of the devices and ALR mechanism. However, they have testify the algorithm with only two pairs of sources and sinks. Rafique et al., (2017) and Fernandez-Fernandez et al., (2016) also provide different energy aware routing solutions based on ALR and turning off inactive links, respectively. However, both works is designed for small networks and not suitable for Large network like GEANT. And if we investigate the heuristics for solving similar kind of optimizing problem without using solver, genetic algorithm is a promising approach as it used by, Galan et al. (2018) to solve the energy aware SDN nodes replacement problem aiming to improve the energy efficiency and Kubler et al. (2012) to improve the connectivity of real time networks. There are also few works which have taken a different approach towards green networking.Gattulli et.al. (2014), Singh and Chandwani (2015) and Zhou et. al (2013) have all tried to achieve green networking in datacentres by using renewable energy. The first and the second one has changed the destination-node based on the availability of the renewable energy and on the third one has considered other criteria like geographical load balancing and server speed scaling. Jo et. al (2018) proposed to choose a cluster head with maximum renewable energy and then choose the member node accordingly. However, one thing is common in above mentioned works is that all of them consider green network same as the energy efficient network. And most of them does not consider doing quality of service (QoS) analysis of their algorithm. They have all tried actively or passively to reduce the energy consumption in order to get an energy efficient network. However, even if energy has a link on sustainability issues, its direct impact on the environment in terms of air pollution and earth’s resource is not explicitly specified. Therefore, in our work, we have introduced two different green metrics CO₂ emission and non-renewable energy usage percentage which address the carbon footprint of the data network more direct. At the same time, we conducted the QoS analysis to measure the applicability of the algorithm.

3. PROPOSED SOLUTION

For our work at first, we have designed an energy model using two green policies: shutting down of the nodes and links if not necessary and adaptive link rate respectively. After that we have proposed our optimizing algorithm which integrates this energy model with the defined objective functions. The number of objective functions is related to number of metrics to optimized. Here we have considered three metrics to test: energy, CEF and NRE.

\[ \varepsilon_v(t) = \int_0^t (\alpha_v + \sum_{w} \delta_{v,w} \beta_{v,w}) e^t \, dt \]  

(1)

As fig. 1 shows, the controller orchestrating the whole system. At first, the controller receives the traffic information from the network, and depending on the geographical position of the nodes, the values of environmental metrics: CO₂ emission and non-renewable energy usage are received. Then, based on the objective function, optimizing algorithm reduces energy consumption or CO₂ emission or non-renewable energy usage respectively and the controller generates a set of data and control plane so the solution plane can be implemented to the network. Data plane provides the information for routing whereas control plane provides the information about the ALR settings. The optimization problem falls into multi-commodity flow class which is known as NP-hard problem as explained by Giroire et. al. (2010). Therefore, in order to solve our optimization problem, we have used the Genetic Algorithm (GA) as our heuristic algorithm. The modelling of the energy equation, the formalization of the problem and the heuristic function are described in the next sub-section.

3.1 Modelling and problem formalization

We have formalized the cost of data and control planes based on the objective function but to do that we needed an energy model and considered a simple energy model proposed by Gupta et.al. (2004). It is assumed here that energy consumption of a node \( v \) follows the linear model:
which are satisfying both throughput demand and the capacity constraint of the links. Even after following the constraints several paths are possible therefore, for each demand set several data planes are possible. Each data plane should respect the following constraint for a demand $i$,

$$w = \Pi_{iw} \text{ such as } \sum_{(keKk \in Ew = w)} \lambda_e \leq C_{vw} \quad (2)$$

Control plane, $\Gamma$, is used to define the controllability of the link capacity that means ALR mechanism. The link capacity is considered zero when both interfaces of a link are turned off which both are multiplicative factor of the energy consumption, the objective function can be rewritten as equation (4) and (5) for CEF and NRE respectively.

$$f(\Gamma) = \arg \min_{\Gamma} \int_0^T \sum_{w \in \Pi} \gamma_w \left( \alpha_v + \sum_{w \in \Pi} \delta_{vw} \beta_v (\Gamma_{vw}) \right) t dt \quad (3)$$

where two binary variables are included to implement the first green policy: turning off network devices when not in use.

$$\delta_{vw}^g = \begin{cases} 1 & \text{if } \sum_{\Pi \in E} \Pi_{iw} > 0 \text{ or } v \in fl \text{ else } 0, \\ 1 & \text{if } \exists k \in K, \Pi_{ik} = w \text{ or } \Pi_{kw} = v \text{ else } 0. \end{cases}$$

Equation (3) serves the purpose when minimizing energy is the objective function. Now if we consider the CEF and NRE, which both are multiplicative factor of the energy consumption, the objective function can be rewritten as equation (4) and (5) for CEF and NRE respectively.

$$f(\Gamma) = \arg \min_{\Gamma} \int_0^T \sum_{w \in \Pi} \Lambda_v \gamma_w \left( \alpha_v + \sum_{w \in \Pi} \delta_{vw} \beta_v (\Gamma_{vw}) \right) t dt \quad (4)$$

$$f(\Gamma) = \arg \min_{\Gamma} \int_0^T \sum_{w \in \Pi} \Psi_v \gamma_w \left( \alpha_v + \sum_{w \in \Pi} \delta_{vw} \beta_v (\Gamma_{vw}) \right) t dt \quad (5)$$

### 3.2 Heuristic Algorithm

We have used genetic algorithm (GA) as our heuristic function. Mutation and crossover approach of GA help to avoid local minima which makes it more robust than any other enumerative solutions. We have considered each data plane as a chromosome. If we have a set of demand, $K = \{1, 2, 3, \ldots, k\}$ and the topology has $E$ number of links, then a chromosome, $\gamma$ consists of $|K| \times |E|$ numbers of 0’s and 1’s. Where 1 represents this link is used for that demand and 0 means that link is not needed for that demand. If one particular link is zero for all the demands, then we can turn off that link. For creating a set of chromosome-pool we have generated a set of data planes a randomized version of depth first search to diversify the result from simple DFS. In the chromosome-pool one solution is always generated by shortest path first (SPF). Therefore, even the initial data plane set provides a solution as same as SPF. Crossover is done with random cut and demand wise. That means, paths for half of the demands are chosen randomly from one parent chromosome and the rest are chosen from another parent and then crossover is done. For mutation, from one chromosome, one demand or more is chosen randomly, and the path for that demand is replaced with a new path generated using random depth first search (RDFS) which traverse randomly in each level in order to achieve more diverse result compare to simple Depth First Search (DFS). The objective functions are used as a fitness functions for GA.

After GA is executed, two things are done consecutively. First, the chromosome is converted into data plane. Secondly, generate the control plane according to the data plane so that controller can make appropriate changes in the topology. The trickiest part of converting chromosome into data plane is, even though each chromosome indicates that if a link is used for a particular demand or not, however, the direction of the flow of the traffic is not defined. For calculating the fitness of the chromosome this information is not necessary as in a bidirectional graph if a link is used in any one direction, it will be considered as in use. But in case of data plane and control plane this piece of a primal information. Algorithm tracks down the path of each demand and generates the data plane accordingly. For generating control plane, the algorithm finds the minimal allocation satisfying the throughput demand gathered by the data plane. So, the controller while applying green policies, shutdowns v nodes when $\delta_{vw}^g = 0$ and output ports w of nodes v when $\delta_{vw}^g = 1$ and decreases the capacities of output ports w of switches v when $\Gamma_{vw} < C_{vw}$.

### 4. PERFORMANCE EVALUATION

#### 4.1 Experiment Setup

For evaluating the performance of the algorithm, we conducted the experiment using GÉANT topology. Fig. 2 provides the current representation of GÉANT network. We wanted to test the algorithm performance with geographically diverse networks that’s why GÉANT is chosen. The topology attributes and traffic are taken from SNDlib (2005). However, size of data traffic has been increased significantly in this last decade. Therefore, traffic has been multiplied by increased
traffic co-efficient extracted from Cisco VNI (2019) and Global IP Traffic Forecast and Methodology, (2012). 25 demands are chosen randomly to conduct the experiment such that there will be a set of nodes which are not involved in planes, hence the shutting down policy may apply. The size of the demands was randomly distributed. For power measurement, values are taken from Addis et. al (2014) and Van et. al (2012). And the values of carbon emission factor and non-renewable energy usage percentage has taken from Brander et al. (2011), IRENA (2019) and Koffi et. al (2017).

For analysing the performance of the algorithm based on quality of service (QoS), the networking metrics: end-to-end delay and packet loss are selected. We have used the network simulator modeler. In the modeler, the average packet size is considered 1500 Bytes. The probability distribution function (PDF) of the packet size is constant. Whereas, PDF of packet inter-arrival time is exponential. We have designed the network on modeler based on our algorithm provided data and control planes.

4.2 Results

Energy, CO₂ and NRE savings comparison:

Fig. 3 shows the comparison of savings of three different criteria for three different objective functions. Results are compared with a green policy enabled SPF. Every metric performs better than the other two when the criteria is same as their objective function. With our EAR with GA (E_GA), we have maximum energy savings compared to the other two approaches. E_GA saves around 24% of energy consumption compared to SPF for this scenario. Whereas CO₂_GA and NRE_GA save around 18% and 23% energy consumption respectively. However, when the minimizing factor is reducing CO₂ emission, CO₂_GA reduce 36% of the CO₂ emission compared to SPF. Whereas NRE_GA and E_GA reduce 31% and 20% respectively. And when the condition is reducing non-renewable energy usage NRE_GA reduces the non-renewable energy usage by 28% compared to SPF. E_GA has outperformed SPF in all three criteria, however, for a green solution the primal focus should be reducing the overall carbon footprint. And in that particular criteria, E_GA is lagging behind from the other two. CO₂_GA reduces almost double of the CO₂ emission compared to E_GA. There is no argument that solutions that are considering environmental metrics consume more energy than E_GA. But if the overall carbon emission of the system is lower by avoiding high CO₂ factor nodes or by avoiding non-renewable power sources, then that means, the system is consuming green energy more than brown energy which should be a primal focus for any green network engineer.

End to end delay comparison:

End to end delay is one of the most crucial QoS parameters to analyse an algorithm. We have compared and analysed the ETE delay ratio with SPF for all metrics by using them as an objective function of the algorithm. It can be observed in a box and whisker plot in fig. 4 to have a better understanding of the distribution of the data.

Usually, SPF has good performance in terms of ETE, therefore, we kept SPF value as a comparison unit for our three algorithms. The graph clearly shows that, NRE_GA has higher ETE delay for packets compare to other two. We observe that, the degradation rate of NRE_GA compared to SPF in terms of average ETE delay of packets are more than 5%. And E_GA has the lowest ETE delay ratio. The placement of nodes in GÉANT network and specially the CEF and NRE values of each node has very high influence in this result. Fig. 5 shows the different topologies for different solutions which are later designed in modeler for QoS measurements. Now, if we see the topology of for NRE_GA, 80% of the overall demands are passing through node-4 (Germany). The main reason behind this kind of bottlenecking is the NRE value of the other nodes.
For example, node-16 (France) uses nuclear power plant for producing energy and because of that, it has a very high NRE value. Similarly, node-2 (Poland) uses coal for producing energy and has also very high NRE value. Therefore, when NRE is considered as an objective function all the paths avoid using node-16 and node-2 which creates a bottleneck in the topology that causes the unexpected ETE delay. For E_GA also, this kind of scenario might occur but for our test case it was not the case.

Secondly, for both CO2_GA and NRE_GA each node has different CEF or NRE values. Which makes the routing more complex than straightforward energy aware routing where all the nodes are considered as equal in terms of energy consumption. For a similar sort of reason, CO2_GA underperforms compared to E_GA in terms of ETE delay.

**Packet loss:**
Fig. 6 shows the packet loss for all three metrics plus the packet loss of SPF. In terms of overall packet loss, all of them perform similarly even though there were few outliers in the data set. All of them have an average of less than 5%. And for at least 75% of the overall demands for all algorithms, the average packet loss percentage is close to zero. The outliers are expected as we have saturated the network with maximum demand size of 97Gbps whereas the maximum amount for bandwidth was 100Gbps. This has been done in order to see the robustness of the algorithm. Packet loss or ETE delay for packets is the result of when the buffer is full or there is a congestion in the link. For both CO2_GA and NRE_GA, packet loss and ETE delay are slightly higher than E_GA and E_GA has slightly higher value than SPF. It is expected as all of them tries to reduce the number of nodes and links to maximize the greening of the network.

Additionally, ALR is maximizing the link utilization percentage and minimizing energy consumption at the same time. For E_GA, the goal is to reduce the number of nodes and links regardless of which node or link, for CO2_GA and NRE_GA the goal is also to shut down nodes and links as both of them are a multiplicative function of energy and shutting down reduce the overall value of CEF and NRE. However, as different node has different green metrics values, it curves the network even more. So, the packet loss is almost similar for all the cases and in worst case scenario (for NRE_GA) we get an average delay ratio of 5% compared to SPF. Whereas, when comes to savings of CO2 emission and energy consumption, the same algorithm can save more than 31% and 21% respectively compared to SPF. Now, if we analyse the trade-off between quality of service and quality in sustainability, our
proposed algorithm clearly shows significant improvement in terms of environmental impact with a slight effect on time performance, offered by the network. Additionally, we can say that, by using our proposed metrics (carbon and renewable energy) in the objective function provides better opportunity to reach the goal of greening network compared to using only energy as objective function which is a common practice in the literature for green networking.

5. CONCLUSIONS

In this paper, we have proposed two new metrics for greening network. The result clearly identifies the need for looking at the problem of greening network beyond from the angle of energy efficiency. In terms of CO₂, when reducing CO₂ is considered as objective function, the same algorithm reduces almost twice of the carbon footprint compared to an energy efficient algorithm which is not negligible. Similarly, when NRE is considered, it also consumes less CO₂ than an energy efficient solution. We have proposed a solution which gives the required control plane and data plane for implementing it on SDN platform. And as a part of QoS analysis ETE delay and packet loss is also measured and evaluated for the algorithm with all three metrics alongside with SPF. It showed that, greening the network is not against with QoS parameters. For future work, we want to conduct temporal analysis for our introduced environmental parameters as during different time period of a day, electricity production technique can vary, and algorithm might provide intriguing results.

REFERENCES