Rethinking Allocation in High-Baseload Systems: A Demand-Proportional Network Electricity Intensity Metric

DANIEL SCHIEN, University of Bristol, Great Britain
PAUL SHABAJEE, University of Bristol, Great Britain
CHRIS PREIST, University of Bristol, Great Britain

Additional Key Words and Phrases: Environmental sustainability, environmental assessment, Internet energy consumption

1 INTRODUCTION

As companies and and entire economic sectors begin to respond to the urgent need to reduce carbon emissions and support global decarbonisation, robust assessments of environmental impact are essential to guide their actions. Network operators are increasingly aware and take action to decarbonise their infrastructure, for example by procuring renewable electric. However, the organisations providing the application services that run over the networks, and for which the networks are built and operated, do not currently know how they can support the reduction of the environmental impact. This lack of alignment in decision making is inefficient and undermines the sector's decarbonisation goals. For high-throughput digital services, such as downloading video games or streaming videos, the current electricity intensity metrics for assessing the energy consumption from use of internet networks allocate a share of energy consumption per average data volume over a duration of typically one year. These metrics thus make no distinction when a service is using the network. However, absolute energy consumption of network devices is strongly dependent on their maximum bandwidth. Increases of peak-time throughput thus drive future expansion of the network capacity, electricity consumption and hardware replacements. By ignoring the variability of demand, electricity intensity metrics disregard the influence that peak-time use of the network has on future environmental impact and overestimate the environmental effects of off-peak use. In this text we investigate the design of an alternative energy intensity metric that redistributes burden of baseline power consumption proportional to data throughput. Such metrics can incentivise demand-shifting of data traffic and thus reduce the pressure on network expansion, which can contribute to a reduction of carbon emissions long-term. We illustrate this approach with an example and consider how it can be combined with carbon intensity metrics for carbon footprinting.

2 BACKGROUND

2.1 Carbon Footprinting Approaches

A set of standardised methods exist to quantitatively estimate environmental impacts due to the life-cycle of a physical good or service.. The standard methodology for analysis of the environmental impacts of a product or service over their lifetime is Life Cycle Assessments (LCA). Besides LCA some informal practices that are loosely associated to streamlined LCA [Gradin and Björklund 2021] can be found in the body of related carbon footprinting studies. While full LCA often consider a range of environmental impact categories besides global warming potential, such as eutrophication or ozone depletion, carbon footprints are exclusively focused on estimating greenhouse gas emissions. Environmental assessments serve a variety of goals, including "informational" (e.g. marketing claims and consumer information) as well as "change-oriented" uses such as product and policy development, design choices, hot spot elimination [Weidema 1998]. The use of carbon footprints to support decarbonisation initiatives would fall into this change-oriented category. In order to enable this goal, carbon footprint methods must provide relevant information that stakeholders require to take action towards reducing carbon emissions.

For the assessment of impact during the use phase carbon footprints of energy-using products (i.e. electricity), or services built using them, usually develop an electricity footprint of the total electricity consumption to which then carbon intensity factors for electrical energy [DEFRA et al. 2014] are applied.

Carbon footprints can be carried out for an organisation or a system in total, such as the entire networking infrastructure devices operated by an ISP, or for individual services using the infrastructure, such as the download of a large file; so-called product carbon footprints. In this text we are particularly interested in product carbon footprints of application services and the data transfer that their constituent connections involve. Examples of product carbon footprints for digital services include [Schien et al. 2021; Williams and Tang 2013].

Networked software services, such as the World Wide Web, video streaming or social media, depend on a product system in which user devices are connected to an electronic supply chain of network devices and servers that consume electricity and are run by separate organisations. Our main concern in this text is to enable pro-environmental decision making by providers of software services and consumers through greater transparency of the relationship between infrastructure use and environment impact. In this analysis of electricity intensity metrics we focus on networking devices in shared networked (ISP networks). However, the principles apply to other types of inelastic energy-using infrastructure - such as servers in datacentres.

2.2 Electricity Intensity Coefficients

An electricity intensity is a coefficient that can be multiplied with a measure of system-use to calculate output electricity consumption. For core networks, mainly data volume has been used as the intensity metric. With this metric the electric energy can be calculated as $E = v \cdot I_v$ with I_v being the energy intensity coefficient per unit of data volume, and v the data volume of the service for which the carbon footprint is being calculated. Note, for access networks, time has been used to model electric energy, as in $E = t \cdot I_t$ [Schien et al. 2013]. Then I_t is equivalent to allocating a share of the power consumption of the infrastructure during the duration t. This could be the time for watching a movie.

The estimation of these coefficient has as long a history as their use in estimating energy footprints of services, with [Koomey and Berkeley 2009] being among the earliest. These intensities have been derived via top down or bottom-up approaches [Schien and Preist 2014a,b]. Roughly, a bottom-up model sums the energy intensity of network components along an average route through a network, e.g. [Baliga et al. 2009]. While top-down approaches estimates the energy consumption of an entire system over a time frame (normally a year) and divide this by the total data volume transported during this time. Recently [Aslan et al. 2018] provide a review of intensity values. Their electricity intensity value is also recommended by the GHG Protocol ICT sector guidance [Carbon Trust 2012] in assessments of services.

Common to all electricity intensity coefficients is that they are scalar values. This approach has been widely criticised [Carbon Trust 2021; Koomey and Masanet 2021; Malmodin 2020] for ignoring the low energy proportionality of networks, a concept usually associated with [Barroso and Hölzle 2007]. While the energy intensity coefficient in their current popular form suggests that the total power consumption P_T of a system is a linear function of the data volume that it transports, in practice power draw of network devices is largely independent from throughput (hence of low energy proportionality). Instead, the power consumption is better understood as the sum of a static baseline power consumption P_b that is independent from use, and a dynamic portion that scales with utilisation P_u, such that $P_T = P_b + u \cdot P_u$. For network devices, baseline power is close to 80% of total power consumption [Chan et al. 2016].

3 COUPLING ELECTRICITY INTENSITY WITH CHANGE OF BASELOAD POWER

This baseload electricity use needs to be addressed as part of the decarbonisation of the ICT sector for at least two reasons:

(1) It contributes directly to carbon emissions through electricity grid emissions
(2) It uses electricity that could be used elsewhere, thus indirectly increasing the overall carbon intensity of generated electricity

Reducing baseload power consumption is a challenge for engineers building networking devices and network architects. However, the network co-evolves with the digital software services that use it [Preist et al. 2016]. The designers and software engineers building the software services need to also be empowered to make decisions that help reduce baseload power consumption. Electricity intensity coefficients in their current form are not providing this change-oriented decision support that is required of carbon footprints.

The current intensity coefficients suggest a direct proportionality between electricity consumption and throughput that does not correspond to the actual behaviour of network device electricity consumption. Any projection of electricity footprints that extrapolate this proportionality by changing data volumes are thus inappropriate. For example [Obringer et al. 2021] apply such reasoning and estimate that network throughput reductions from turning off video in video calls, would reduce monthly carbon emissions from 9.4 kgCO2e to 377 gCO2e for a typical individual. However, due to the lack of energy proportionality, changes of data throughput do not result in substantial changes in electricity consumption in the near term. This can be illustrated by the absence of increases of electricity consumption during Covid - despite significant increases in data throughput [GSMA 2020].

In response to this, [Malmodin 2020] propose a model that separates a constant, per-subscriber portion of inelastic baseline power draw P_b that is combined with service use time, and calculate an intensity metric for only the dynamic power consumption I_v^d that scales with data volume in: $E = P_b \cdot t + v \cdot I_v^d$. This approach offers a better representation of the (small) instantaneous change in energy consumption from a change in transported data volume. However, that model does not improve on the more substantial problem of baseload power consumption. By making baseload independent from data volumes, the model in fact decouples the use of the system from its electricity consumption. This is clearly not representing the the effect that demand for network bandwidth exerts on the evolution of networks. 1 illustrates how demand drives capacity and in turn power consumption. As the throughput from services increases, additional bandwidth (capacity) is provided by replacing network nodes. After upgrades, the baseline power consumption steps up. The dynamic power consumption contributes to a relatively much smaller degree to overall power consumption.

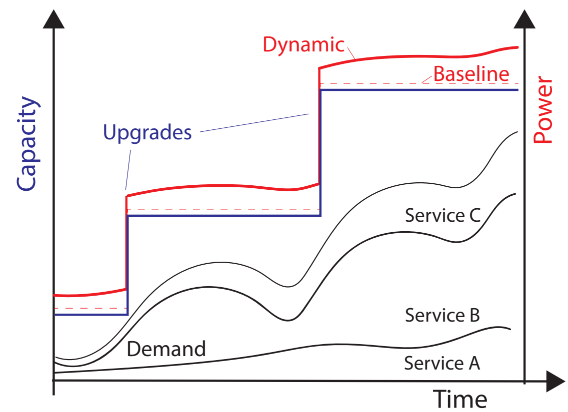

Fig. 1. Scale-free schematic of capacity (blue), power consumption (red) and demand (black) over time. As the throughput from services increases, additional bandwidth (capacity) is provided by replacing network nodes. After upgrades, the baseline power consumption steps up. The dynamic power consumption contributes to a relatively much smaller degree to overall power consumption.

To meet the need for carbon footprints to provide change-oriented decision support we need to identify the key casual factors that determine the change in the baseload power draw and explore how those factors can be fed back into the impact metric - in this case the energy and carbon footprints. We argue that networks could have a lower power consumption at off-peak periods (i.e. baseload), if the peak demand was lower. Electric energy consumption of network devices is defined by their peak throughput. Notwithstanding background efficiency improvements, and assuming similar functionality, a network devices with lower peak capacity will tend to have a lower power consumption. It is thus the maximum capacity that determines the device class and it's base and dynamic energy consumption. And the choice of network device thus depends on the peak capacity the network is designed to provide.
In order to address the lack of causal connection between use of service at peak and baseline power consumption, we present a transformation of the electricity intensity metric. This updated intensity metric has the potential to incentivise alternative system design and behaviour change, leading towards a lower growth in peak and baseline power consumption.

4 A TRANSFORM FUNCTION FOR A CHANGE-ORIENTED INTENSITY METRIC

The transformation burdens data traffic at peak time with proportionally higher share of the baseline power consumption than traffic at other times. For this, we consider patterns of variability of demand as our starting point. For this current investigation we use a representative shape of diurnal demand.
We model the electricity consumption of the network as the sum of the baseload (E_b; straightforward to measure during idle periods) and the dynamic remainder due to utilisation (E_u). This information is available to the network operator. Total energy consumption is then given by $E_T = E_b + E_u$ and for each 30-minute interval i of metered electricity consumption (i) this is: $E_{T_i} = E_{b_i} + E_{u_i}$
Given a pattern of demand with peak throughput V_P, we then reallocate baseload energy consumption. For this we define a transform function C_i that 1) scales the data volume in each 30-minute time window inverse proportionally to peak traffic and then re-normalises, so that the overall volume remains constant.

$$C_i = \frac{(\frac{V_i}{V_P}) \cdot V_i}{\sum_{i=1}^{48}[(\frac{V_i}{V_P}) \cdot V_i]}$$

This transform function is then applied to scale the baseline energy E_{b_i} with C_i in each 30-min interval as $E'_{b_i} = E_{b_i} \cdot C_i$. We then add the adjusted baseline power to the dynamic power as $E'_{T_i} = E'_{b_i} + E_{u_i}$. Finally, we divide the adjusted total energy consumption by the data volume per half-hourly interval to calculate the new intensity metric V'_i as $I'_i = E'_{T_i}/V_i$

In Figure 2 we show the demand-proportional intensity metric relative to the constant electricity intensity metric.

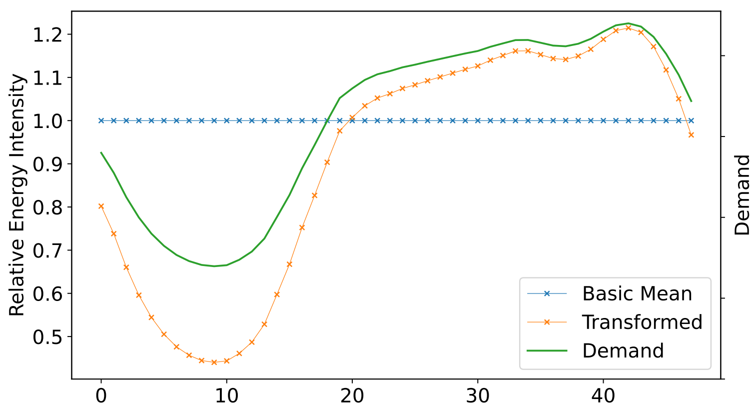

Fig. 2. Network Electricity Intensity Metric over 48 half-hourly intervals. Basic Mean (current practice), Transformed (reallocated baseline energy). Demand - Typical diurnal traffic shape for core networks. From London Internet Exchange (LINX) https://datasciencecampus.ons.gov.uk/projects/what-can-internet-use-tell-us-about-our-society-and-the-economy/. Note, this is only for illustration, core traffic will be likely to follow slightly different patterns than domestic environments, in particular during working hours.

The result of this simple transform is that each unit of activity in any given period is weighed such that it is largest at periods of peak demand and lowest at the lowest period. Other kinds of transform might shift the intensity differently, e.g. a steeper function would weight peak activty even more highly.

4.1 Demand Proportional Carbon Emissions

It should be noted that here we are simply looking at energy use, in order to extend the analysis on GHG emissions we could either use the results of the energy transform function and calculate the emissions intensity in each 30-minute interval from the re-allocated energy intensity values in each period. Alternatively we could calculate the total emissions in each 30 minute interval prior to the transform (mean kgCO2e/kWh intensity varies over the day) and use the transform function to re-allocate the GHG emissions as we have the electrical energy.

Taking this yet one step further, we could use dynamic carbon intensity, i.e. mean hourly carbon intensity (gCO2e/kWh by hour of the day). Such data is now readily available for many grids, e.g. the UK National Grid. In Figure 3 we compare the resulting carbon intensity metrics for data when combined with a dynamic carbon intensity. In this case the carbon intensity is the 30-minute average over the year 2021. The carbon intensity of electricity varies significantly, thus amplifying the effects of the transform function.

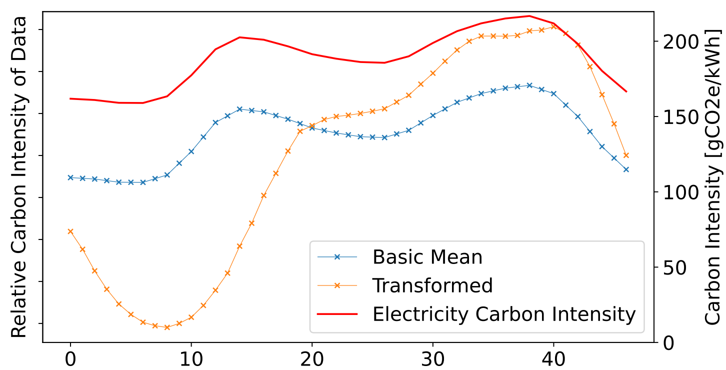

Fig. 3. Comparison of Data Demand Carbon Intensity Metrics. Carbon intensity for electricity as the 30-minute average of the UK National Grid in 2021. This electricity carbon intensity is combined with basic mean electricity intensity of data and the transformed electricity intensity, respectively.

5 DISCUSSION

A near ubiquitous feature of electrical energy using components in digital service system is that they have a baseload of energy demand. These include individual, low level, components such as internet routers, servers and end user devices and also large scale aggregate components including data-centres which are composed of servers, networking equipment, data storage devices, building infrastructures and services, etc. By baseload we mean that there is a minimum of energy use regardless of demand or utilisation of a component. Within carbon accounting, such baseline power consumption forms an overhead that needs to be allocated to services in the product system. Note - embodied GHG emissions (from raw material extraction, manufacturing and transport) as well as disposal form similar overheads.

As we have illustrated, the existing electricity intensity metrics used for carbon footprints of digital services do not support the stakeholders in taking causally informed action during design and use of digital services that can support reducing peak capacity growth and thus support decarbonisation. Our very simplistic example illustrates how such signals and metrics could be developed in GHG impact allocation. We present an approach, rather than a finalised method. More work is required to understand the causal mechanisms that drive network capacity growth.

Creating metrics that are more causally related to infrastructure overheads is not at all new. Such approaches are widely used, for example, so called 'use-of-system' charges used in electricity grids to incentivise behaviours that reduce use of electricity at times of peak demand with the explicit goal of constraining growth in the daily and annual peak demand and so reducing the need for unnecessary infrastructure development and reducing costs (and also carbon emissions). For example, in the UK grid transmission system, the TNUoS charge system consider the three top peak 30-minute intervals in a given year and distributes high charges to consumers during these periods. Such an approach has various side-effects. Comparing this to our proposed transform function, the equivalent for data would be that baseline power is allocated exclusively to the services that use at peak.

The electricity intensity metric we illustrate incentivises demand-shifting. This is a well-established activity that supports the decarbonisation of the electricity sector; for example with variable time pricing.

Our illustrative metric can be combined with a time-of-use metric as presented by [Malmodin 2020] by applying the illustrated transform to the constant base power consumption and adding a term to more accurately model dynamic energy use:

$$E_i = I'_{b_i} \cdot v_i + v_i \cdot I_v^d$$

The choice of time-window that is considered during the analysis affects the design of the metric. We chosen to consider diurnal variability. However, those diurnal peaks vary steadily increase from background increases in consumption. Over the course of a year there are also seasonal patterns that have their own peaks. And thirdly, peak demand is increasingly observed during live sport events and game releases; that exacerbate high demand from other more consistent demand such as that of video streaming.

The foundation for the practical application of such metrics would be a regular publication of a transformed electricity (or carbon) intensity metric by ISPs, which would then be applied to time-of-use data on service consumption. This would enable service providers (e.g. media companies) to integrate the carbon effects from varying time-of-use into their decisions. To consumers understanding of variable carbon intensities could increase acceptance for design interventions.

5.1 Conclusions

In this text we investigate the design of an electricity intensity metric with the goal to include the effect of peak-time data demand on baseline power consumption. We apply a transform function that re-allocates baseline energy towards peak periods and illustrate it’s effect on carbon intensity. We suggest that the principle might be further explored to help the ICT sector more effectively reason about what drives its carbon emissions and how to reduce them.

REFERENCES

Joshua Aslan, Kieren Mayers, Jonathan G. Koomey, and Chris France. 2018. Electricity intensity of internet data transmission untangling the estimates. *Journal of Industrial Ecology* 22, 4 (2018), 785–798. https://doi.org/10.1111/jiec.12630

J. Baliga, R. Ayre, K. Hinton, W.V. Sorin, and R.S. Tucker. 2009. Energy Consumption in Optical IP Networks. *Journal of Lightwave Technology* 27, 13 (7 2009), 2391–2403. https://doi.org/10.1109/JLT.2008.2010142

Luiz André Barroso and Urs Hölzle. 2007. The Case for Energy-Proportional Computing. *IEEE Computer* 40, 12 (12 2007), 33–37. https://doi.org/10.1109/MC.2007.443

Carbon Trust. 2012. GHG Protocol Product Life Cycle Accounting and Reporting Standard ICT Sector Guidance. https://ghgprotocol.org/sites/default/files/GHGP-ICTSG-ALLChapters.pdf

Carbon Trust. 2021. *Carbon impact of video streaming*. Technical Report. https://prod-drupal-files.storage.googleapis.com/documents/resource/public/Carbon-impact-of-video-streaming.pdf

Chien Aun Chan, André F. Gygax, Christopher Leckie, Elaine Wong, Ampalavanapillai Nirmalathas, and Kerry Hinton. 2016. Telecommunications energy and greenhouse gas emissions management for future network growth. *Applied Energy* 166 (2016), 174–185. https://doi.org/10.1016/j.apenergy.2016.01.007

DEFRA, RICARDO-AEA, and carbon smart. 2014. Greenhouse Gas Conversion Factor Repository. http://www.ukconversionfactorscarbonsmart.co.uk/

Katja Tasala Gradin and Anna Björklund. 2021. The common understanding of simplification approaches in published LCA studies—a review and mapping. *International Journal of Life Cycle Assessment* 26, 1 (2021), 50–63. https://doi.org/10.1007/s11367-020-01843-4

GSMA. 2020. COVID-19 Network Traffic Surge Isn’t Impacting Environment Confirm Telecom Operators. https://www.gsma.com/gsmaeurope/latest-news-2/covid-19-network-traffic-surge-isnt-impacting-environment-confirm-telecom-operators/

Jonathan Koomey and Eric Masanet. 2021. Does not compute : Avoiding pitfalls assessing the Internet’s energy and carbon impacts. *Joule* (2021), 1–4. https://doi.org/10.1016/j.joule.2021.05.007

Jonathan G. Koomey and Lawrence Berkeley. 2009. Whole system redesign for electricity efficiency: Some lessons from the world of cloud computing. (2009).

Jens Malmodin. 2020. The power consumption of mobile and fixed network data services - The case of streaming video and downloading large files. *Proceedings of the Electronics Goes Green 2020+ Conference, Berlin, September 1* (2020), 87–96.

Renee Obringer, Benjamin Rachunok, Debora Maia-Silva, Maryam Arbabzadeh, Roshanak Nateghi, and Kaveh Madani. 2021. The overlooked environmental footprint of increasing Internet use. *Resources, Conservation and Recycling* 167, January (2021), 105389. https://doi.org/10.1016/j.resconrec.2020.105389

Chris Preist, Daniel Schien, and Eli Blevis. 2016. Understanding and Mitigating the Effects of Device and Cloud Service Design Decisions on the Environmental Footprint of Digital Infrastructure. In *CHI'16 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, San Jose, 1324–1337. https://doi.org/10.1145/2858036.2858378

Daniel Schien and Chris Preist. 2014a. A Review of Top-Down Models of Internet Network Energy Intensity. In *ICT for Sustainability (ICT4S)*. Stockholm, 1–8.

Dan Schien and Chris Preist. 2014b. Approaches to energy intensity of the internet. *IEEE Communications Magazine* 52, 11 (11 2014), 130–137. https://doi.org/10.1109/MCOM.2014.6957153

Daniel Schien, Paul Shabajee, Jigna Chandaria, Dan Williams, and Chris Preist. 2021. Using behavioural data to assess the environmental impact of electricity consumption of alternate television service distribution platforms. *Environmental Impact Assessment Review* 91, September (2021), 106661. https://doi.org/10.1016/j.eiar.2021.106661

Daniel Schien, Paul Shabajee, Mike Yearworth, and Chris Preist. 2013. Modeling and Assessing Variability in Energy Consumption During the Use Stage of Online Multimedia Services. *Journal of Industrial Ecology* 17, 6 (12 2013), 800–813. https://doi.org/10.1111/jiec.12065

Bo Weidema. 1998. Application typologies for life cycle assessment. *International Journal of Life Cycle Assessment* 3, 4 (1998), 237–240. https://doi.org/10.1007/BF02977574

DR Williams and Yinshan Tang. 2013. The Impact of Office Productivity Cloud Computing on Greenhouse Gas Emissions. *Environmental Science & Technology* 47, 9 (2013), 4333 – 4340. http://henleymba.com.hk/web/FILES/tsbe/Williams_TSBE_Conference_Paper_2012.pdf