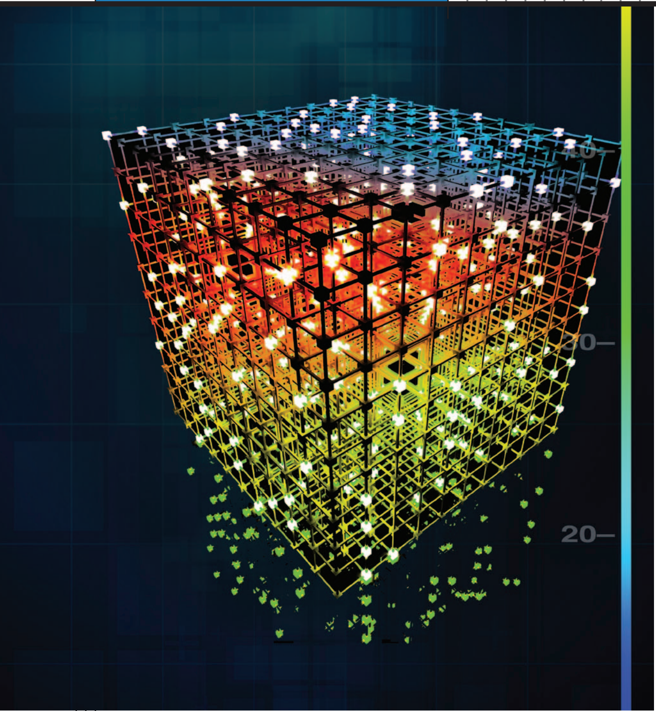


Fundamentals of Green Communications and Computing: Modeling and Simulation

Murat Kocaoglu, Derya Malak, and Ozgur B. Akan, *Koc University, Turkey*

A layered architecture incorporates the concept of minimum energy consumption for communication links and computer networks with multiple terminals, where emission-reduction approaches based on information theory are impractical.

As society becomes more aware of how carbon dioxide (CO_2) emissions are affecting the environment, the information and communications technology (ICT) community is investigating how to ensure that communication systems consume less energy and thus collectively have a substantially smaller carbon footprint. To that end, researchers have been working on novel techniques to reduce the energy dissipation of point-to-point communication links and computer networks.

Although modifying existing ICT systems can lead to incremental energy savings, the primary purpose of lowering energy dissipation is to reduce the CO_2 emission rate. The relationship between these two ideas is more complex than many organizations realize. The ICT community has long assumed that energy savings will reduce CO_2 emissions,[1] but the amounts of emitted CO_2 and dissipated energy are *not* always linearly related. In some scenarios, optimum CO_2 savings in fact do not stem from the lowest possible energy consumption. Minimizing the total energy consumption of links over all possible network routes does not minimize CO_2 emissions because different components use different energy sources.

A bottom-up approach to minimizing energy dissipation starts with establishing a fundamental energy dissipation limit, which shows the true gap between current and optimum energy savings and which can serve as a springboard for developing techniques to achieve that limit. Unfortunately, for point-to-point communications, such an approach applies only in certain scenarios; using information-theory approaches to derive energy dissipation limits is largely impractical for large-scale networks, and even some small networks, because incorporating the dynamics of practical networks from an information-theory perspective is difficult.

As an alternative way to find the minimum energy consumption in larger networks, we propose combining an architecture that uses Internet layers with a physical layer operating at the fundamental limit of energy dissipation. We can then view energy consumption as the fundamental energy consumed per information bit. For large networks, this layered approach is more feasible than an analysis based on information theory.

Our exploration of this approach highlighted some novel implementation challenges, particularly how to address the effects of constraints on the upper layers—constraints that stem from requiring minimal energy consumption at the physical layer. It also revealed new directions and open issues for green network simulation and standardization.

Published by the IEEE Computer Society

DERIVING FUNDAMENTAL LIMITS

As the "Green Communication Networks" sidebar describes, ICT researchers have many ideas about how to reduce a network's energy dissipation. Significantly reducing the carbon footprint of ICT systems is not easy, particularly if the solution is to use only energy-efficient techniques and algorithms to improve existing network components. For the past decade, ICT researchers have actively sought to satisfy the demand for high performance and high bit rates, optimizing communication links and networks to maximize data rate within power and energy constraints.

To continue meeting demand and satisfy the need for lower energy consumption, the ICT community must take a more revolutionary approach to developing green ICT solutions: optimizing ICT systems with the objective of minimizing energy consumption while using data rate and reliability as constraints could significantly reduce ICT's adverse effects on Earth's atmosphere. A cornerstone of this new order is the creation of fundamental energy limits that are independent of a particular technology.

Energy limits

Recently, ICT developers have become more interested in designs that reduce energy dissipation, a focus that begs two questions: is there a fundamental limit for minimum energy per bit independent of technology, and, if so, can that limit be used as the basis for deriving a fundamental CO_2 emissions limit? Pointing out the fundamental limits and revealing the maximum energy savings can aid in developing greener technology by helping to close the gap between current and optimal operational costs and efficiency.

Researchers have explored the energy minimization problem from many perspectives. Studies generally have distinct starting points, but end up with common findings. The main concern is energy per bit or power per rate.

The fundamental energy limit problem considers the transmission of one information bit in a point-to-point communication link. Claude E. Shannon, the father of information theory, analyzed point-to-point communication by mathematically defining *information-theoretic channel capacity*—the maximum achievable rate per channel use. On the basis of Shannon's findings, the minimum energy required for the reliable transmission of one information bit in a point-to-point communication link is $k_B T \ln 2$, where k_B is Boltzmann's constant and T is absolute temperature. (Boltzmann's constant is the physical constant relating energy at the individual particle level with temperature.)

From the computing perspective, one information bit (which is a two-state system) requires $k_B T \ln 2$ amount of energy, obtained from applying Boltzmann's formula. This relationship supports the claim that communications and computing are essentially the same processes, except that observers have different frames of reference.

Consequently, we can assume that the fundamental limits of communications and computing could be the same, although practical considerations such as energy loss during transmission would reflect differences in daily use. Therefore, ignoring the practical issues and focusing on the theoretical limits, $k_B T \ln 2$ is a convenient metric for defining the fundamental energy limit of communications and computing.

One research group used the energy per bit limit to define the *absolute energy efficiency metric*,[2] given as

$$dB\varepsilon = 10\log_{10}\left(\frac{Energy\,/\,bit}{k_B T \ln 2}\right).$$

Figure 1 shows the absolute energy efficiency of state-of-the-art technology. Networks are the least energy-efficient, while a single photon is the most efficient.

Although $k_B T \ln 2$ is a convenient physical metric for the fundamental energy limit of communications and

GREEN COMMUNICATION NETWORKS

ICT service providers have always been primarily concerned with meeting consumers' capacity and quality-of-service demands. However, the recent boost in energy costs in parallel with the increased range of applications and always-on devices are underlining the ICT industry's environmental impact, causing concern about communication systems' collective carbon footprint. Analysts expect ICT systems to account for three percent of the total carbon footprint by 2020.[1]

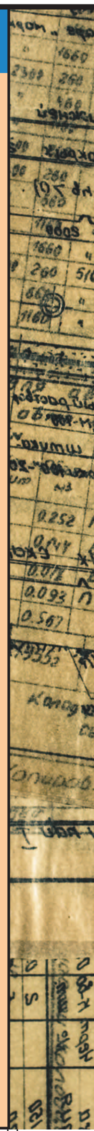
Green communications can mitigate the excessive energy dissipation and achieve sustainable energy-efficient communications. Energy dissipation implies CO_2 emissions via fossil fuel-based energy sources. Therefore, green communication technologies are imperative to alleviate global warming and climate change problems.

Today, the green communications concept focuses mainly on developing energy-efficient communication techniques for networks. Three main approaches are suggested for power management in communication networks:[2] do less work, reduce operating speed, and turn off idle elements. Doing less work means optimizing processes so that the system executes fewer operations and thus uses less energy. Decreasing operation speed could prevent redundant resource use from the mismatched speed of subprocesses.[2] Finally, shutting down idle network components and links can obviously reduce energy dissipation.

Other strategies for reducing communication systems' carbon footprint are to use renewable energy, biodiesel, and solar- and fuel-powered cell sites, and to install fuel catalysts and free cooling units.[1] Using renewable energy sources will decrease the carbon footprint without necessarily minimizing energy consumption.

References

1. W. Vereecken et al., "Overall ICT Footprint and Green Communication Technologies," *Proc. 4th Int'l Symp. Comm., Control, and Signal Processing* (ISCCP 10). IEEE, 2010; http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5463327.
2. M.D.S. Bhawan and J.L.N. Marg, "Recommendations on Approach Towards Green Telecommunications," 2011; http://trai.gov.in/WriteReadData/recommendation/documents/green_telecom-12.04.2011.pdf.



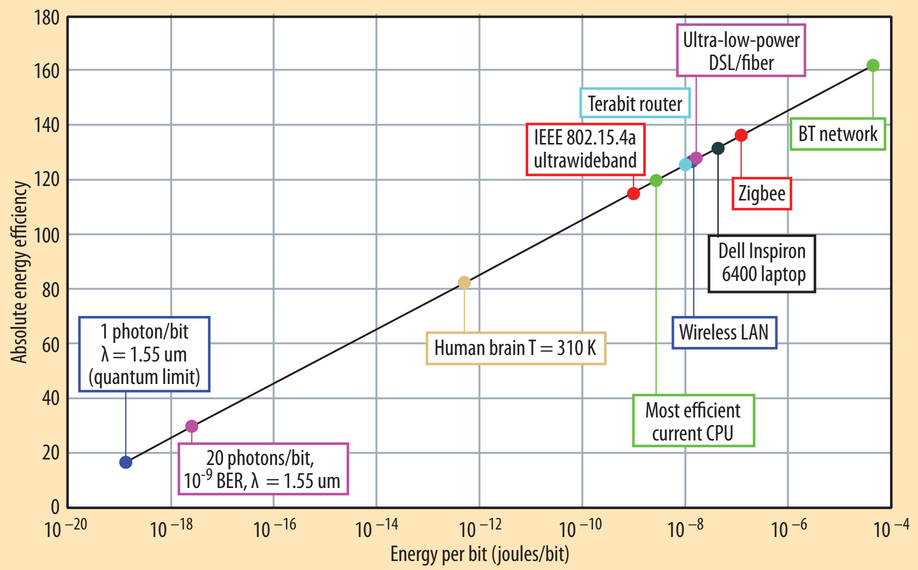


Figure 1. Energy efficiency of state of the art and various limits. The most efficient CPU is still much less energy-efficient than the human brain and far from the efficiency of a single photon, which represents communication's physical limits. T: absolute temperature in Kelvin; BT: British Telecom; BER: bit error rate. (Figure from M. Parker and S. Walker, "Roadmapping ICT: An Absolute Energy Efficiency Metric," *IEEE/OSA J. Optical Comm. Networks*, Aug. 2011, pp. A49-A58.)

computation of one information bit, it is not clear if energy dissipation is required for transmission of one bit. In 1987, Rolf Landauer, an IBM physicist, claimed that $k_B T \ln 2$ is the energy that the system must invest to transmit one bit; it is not the energy dissipated.[3] Landauer also proposed that dissipation is required only if information is lost, which corresponds to irreversible logical operations, such as erasure.[3] If a computing or communication device could use only operations that are logically reversible, he reasoned, the dissipated energy per bit could be arbitrarily small. The ICT community has hotly debated this idea, since Landauer's theory advocates having no energy limit in communications.

In a letter recently published in *Nature*, researchers verified Landauer's vision that $k_B T \ln 2$ is the energy required to erase one bit.[4] These researchers represented one information bit as a single particle's position, which they switched by changing the potential energy function, thereby trapping the particle. They then ran an erasure protocol to reset the bit by moving the particle, or bead, to a predefined side. Their measurements show that, as the protocol runtime spans a larger interval (the bead's motion is slower), the heat dissipation approaches Landauer's limit. Most researchers are now predicting that computers operating at Landauer's limit will be possible in a few decades, yielding energy savings several orders of magnitude greater than in current computers.

Although the $k_B T \ln 2$ limit seems the most relevant bound on minimum energy per bit for current ICT systems, different limits could arise in different settings and problem definitions. Shannon founded his channel capacity derivations on the idea of maximizing information between transmitter and receiver. Analysts can easily derive Landauer's $k_B T \ln 2$ bound as the limiting value of the power-per-information-rate function using Shannon's channel capacity (maximum information among all input distributions).

A more direct investigation could aim to minimize the power-per-information-rate expression, rather than power per channel capacity. This problem is different from minimizing the channel capacity's multiplicative inverse, since the power is in the objective function, rather than being a constraint. Thus, we can find the information rate per channel use, where optimum consumption occurs when the energy per bit attains a minimum value. Such an optimization problem can be defined as

$$\min_{p_x} \frac{\sigma_x^2}{I(X;\,Y)},$$

where σ_x^2 is the signal power at the receiver and $I(X;\,Y)$ is the mutual information—the achievable data rate per channel use between transmitter and receiver. Minimization is over p_x, which is the source probability distribution set. It is possible to impose additional requirements such as satisfying certain data rates or a maximum bit error probability on such optimization problems to ensure that users receive a certain service quality.

CO_2 limits

Traditionally, researchers aim to reduce the carbon footprint by reducing energy consumption, but the implicit assumption in this strategy is that the emitted amount of CO_2 is a linear function of the dissipated energy. Even though no finding suggests a nonlinear relation between the two, minimizing CO_2 as the main objective rather than reducing energy consumption could bring new constraints.

Power plant efficiency, for example, is an important consideration in reducing the carbon footprint. In thermal power plants, efficiency depends on the ratio of average to peak load power, so if the power from the generator is fixed over time, the same amount of energy will yield less CO_2 emissions.

For ICT systems, an open issue is how to develop algorithms and techniques to minimize CO_2 emissions but keep power dissipation almost constant. For data modulation at the physical layer, for example, it is preferable to use binary phase-shift keying, rather than on-off keying, or to use additional precautions to satisfy the fixed power constraint.

One technique for reducing the carbon footprint[1] is to use energy sources with zero CO_2 emissions. In general, the conversion factor between energy dissipation and CO_2 emission depends on the energy source. Renewable sources such as wind or hydroelectric energy emit no CO_2 during operation, and only minute amounts overall, even considering their long life cycles. However, it seems impractical to replace all energy sources with renewable energy in a short time. An alternative is to position the network components with the highest energy dissipation close to green energy sources, such as wind turbines or solar energy fields.

Table 1 lists common energy sources with their emission factors.[5] Given these energy equivalents, calculating the total carbon footprint is straightforward.

Even though researchers have exhaustively attempted to determine the existence of a fundamental energy limit, no one has yet considered the fundamental limit of CO_2 for communications. Work on the fundamental energy limit has promoted interest in green ICT solutions, but we believe that understanding the fundamental limit of CO_2 minimization will lead to solutions that are optimal for the environment.

Table 1. Energy sources, their life cycle emission factors (CO_2 equivalent), and cost per kilowatt-hour (kWh).

Fuel or resource	Emission factor (grams of CO_2 equivalent per kWh)	Total cost per kWh
Industrial coal	1,050	$0.045
Solar power	13	$0.22
Geothermal power	38	--
Hydroelectricity	13	$0.03
Natural gas	443	$0.10
Heavy oil	778	$0.10
Photovoltaics	32	$0.04
Nuclear	66	$0.04
Wind power	9	$0.08

EXTENDING FUNDAMENTAL LIMITS

Defining a generic network's fundamental energy limits using information theory is impossible if systems have finite bandwidth (or delay) constraints. As an alternative, we can analyze minimum energy per bit from a wider perspective by adding the requirements that stem from the finite nature of physical systems and the communication channel, as well as the constraints that the network's upper layers impose. An open problem in obtaining the fundamental limits of our layered architecture is how to include the delay and bit error rate requirements that the application layer imposes.

In networks with two terminals, we can determine the minimum energy per bit from a low signal-to-noise-ratio approximation of Shannon's capacity expression, which yields the $k_B T \ln 2$ limit. Extending this approach to multiple-access channels, broadcast channels, and relay channels has been the first step toward attaining the minimum energy per bit for a general point-to-point energy bound for multiterminal networks. The minimum energy per bit is known for the Gaussian multiple-access, broadcast, and interference channels,[6] but it is still unknown for the three-terminal Gaussian relay channel, despite some progress.

To calculate the capacity per unit cost in stationary two-user and multiterminal channels with no memory, Sergio Verdu and colleagues[7] modeled channel input using a cost function that associates a nonnegative number with each element of the input alphabet. They then demonstrated the tradeoff between the number of symbols and the cost to send every information unit through the channel.

Most researchers do not treat CO_2 and energy minimizations as separate problems, even though the power generator's efficiency directly affects the carbon footprint. In this regard, any solution to minimize CO_2 emissions in multiterminal networks should maximize generator efficiency. One way to achieve the highest generator efficiency is to keep the power from the plant constant over time by deploying smart power distribution algorithms that stress cooperation among terminals. A centralized computing unit is inadequate to determine the best way to allocate and schedule power to achieve optimum CO_2 emissions.

Distributed computing among network nodes would satisfy a given power or CO_2 emission requirement, but it might also impose problems related to synchronization, parallel computing, and additional feedback requirements. Further, promoting cooperation in expanding large-scale networks could increase delays because obtaining information from distant nodes becomes more difficult as the network grows. Therefore, network delay tolerance should determine the allocation of resources in distributed systems.

Concurrent communication over the same medium now occurs primarily among more than two parties, and networks face many challenges inherent in point-to-point communications. Concurrent and fair use of finite resources, addressing the desired destination with proper relay selection, and flow control to provide matching data

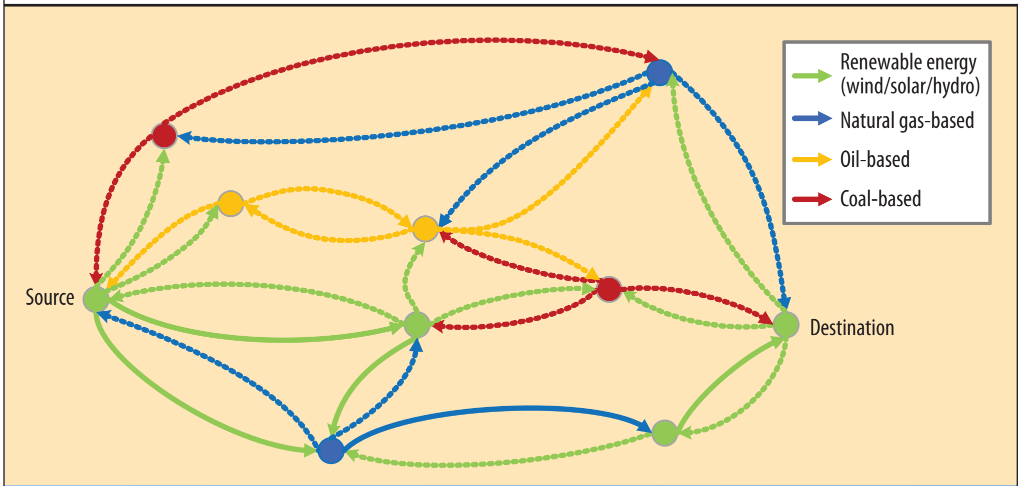


Figure 2. Minimum CO_2 path selection in a network. Assigning carbon cost to each node and its outgoing links produces a directed graph. Solid arrows depict paths that minimize CO_2; dotted arrows depict other paths. The colored arrows in the key box are in order of least CO_2 emissions, with renewable energy generating the lowest amount and coal-based sources generating the highest amount.

rates between transmitter and receiver are only a few of these problems. The idea of separating these problems to provide independent solutions is what led to the Internet's layered structure, although it might not be the optimal solution from an energy efficiency view.

Minimizing energy dissipation

Rather than minimizing energy at each layer separately, our proposed architecture builds the well-known Internet layers on top of a physical layer with minimum energy dissipation per bit and attempts to incorporate CO_2 emission limits into the system design. We encountered several open issues when we considered the different network layers.

In the *media access control* layer, handling channel access with minimum energy consumption is still an open problem. At present, to ensure fair channel access among multiple users, most networks rely on deterministic channel-allocation schemes (for example, time division multiple access) or random access techniques, such as carrier sense multiple access (CSMA) schemes. A network with many users cannot allocate an orthogonal resource set to each user but requires the user to first listen to determine if the channel is busy. CSMA uses carrier detection for this task. Detecting the channel state in a communication system operating at minimal energy limits also presents new challenges.

In the *network* layer, open issues include path determination, logical addressing, and choosing the optimal path when the physical layer is operating at the minimum energy limit. Additionally, the address length determines the amount of information to be processed, increasing the total energy consumption. In the *transport* layer, whether or not flow control uses feedback determines the energy dissipation.

Minimizing CO_2 emissions

Figure 2 shows our routing algorithm that achieves minimum carbon emissions in the network layer. If the delay condition that the application requires is satisfied, the system could use longer multihop paths (instead of the shorter paths) if they have lower emissions. Existing cost minimization algorithms could also help realize the most CO_2-efficient routing scenario.

Future work could focus on developing a CO_2-aware rate adaptation policy in the transport layer, similar to how congestion control policies adaptively adjust each incoming link's rate to prevent congestion. Such a policy could minimize each dataflow's CO_2 emissions by adaptively changing the rate of each node's outgoing links that contribute to the flow.

GREEN NETWORK SIMULATION

The increasing number of network terminals and demand for high performance are driving network energy consumption upward, making energy efficiency a critical part of evaluating network performance. However, extended coverage and the large user population make it extremely difficult to analyze the energy cost of large-scale networks. Consequently, simulation has become an invaluable aid in evaluating network architectures and protocols.

Some simulator designs focus on the network's energy consumption. For example, Hermes[8] is a high-level energy-aware reservation model for end-to-end networks that adaptively manages bandwidth reservations. To save energy, Hermes coordinates the communication infrastructure by regulating traffic flows and putting the infrastructure into sleep state predictively.

Hermes' designers validated the model using the Python-based Bookable Network Simulator, which generates the network topology and bandwidth reservation according to input characteristics. Using this traffic and topology, it then simulates five scheduling algorithms and compares the performance and energy consumption of each. The scheme that yields the minimum estimated energy

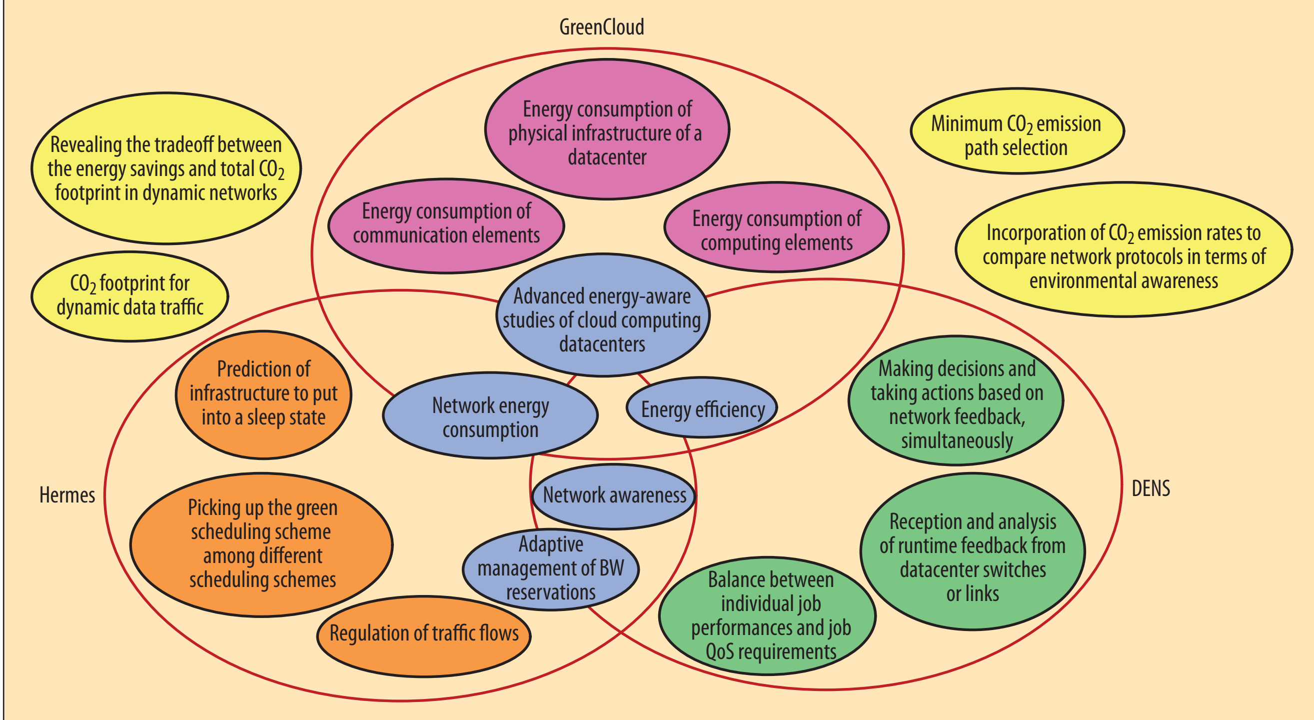


Figure 3. Classification of the Hermes, GreenCloud, and DENS green simulation environments. The blue ovals show core specifications that all three simulators meet. The yellow ovals list what future simulator designs must address to fulfill the aims of green simulation.

consumption for each possible allocation is the best, or greenest, scheduling scheme.

Developers of the GreenCloud simulation environment[9] aimed to extend the packet-level network simulator ns2 for use in advanced energy-aware studies of cloud computing datacenters. Unlike cloud computing simulators such as CloudSim or MDCSim, GreenCloud extracts, aggregates, and categorizes information about the energy that computing, communication elements, and the datacenter's physical infrastructure consume.

Datacenter energy-efficient network-aware scheduling (DENS) combines energy efficiency and network awareness to balance performance and quality-of-service requirements for individual jobs.[9] Network awareness is the ability of DENS to receive and analyze runtime feedback from the datacenter switches and links. DENS can also use network feedback to simultaneously make decisions and take actions.

As Figure 3 shows, state-of-the-art green simulators are concerned primarily with predicting the best tradeoffs between energy savings and traffic demand. The figure also shows what these simulators lack to incorporate the total carbon footprint for dynamic data traffic. For example, the ability to select the minimum CO_2 emission path will reveal the tradeoff between energy savings and total carbon footprint in dynamic networks. Future green network simulators should incorporate CO_2 emission rates so that service providers can compare the environmental impact of network protocols.

GREEN COMMUNICATION STANDARDS

The ICT community needs legitimate policies and standards so that organizations and manufacturers can certify their communication systems as green. Some international approaches assess the efficiency and carbon footprint of telecommunication systems, and standards for industry and communications equipment.[10] Life cycle assessment (LCA), for example, estimates the total environmental impact of products throughout their lifespan by measuring greenhouse gases and determining their grams of CO_2 equivalents. Hence, LCA determines the product's total CO_2 emissions index from its origin through its manufacturing, distribution, use, and disposal.

The Energy Consumption Rating assessment tool determines the ratio of maximum power consumption in watts to effective system throughput in bits per second—essentially an energy versus performance assessment. The tool is also suitable for measuring variable throughput, and there are additional tests for packet-based systems. Comparing product metrics will enable service providers to add energy efficiency to their purchase criteria.

The European Telecommunications Standards Institute has recently announced the second edition of TS 102 706, a technical specification for measuring the energy efficiency of wireless access equipment.[11] The energy efficiency of access networks is critical, since most of the energy in the Global System for Mobile Communications (GSM) dissipates in access networks. The TS 102 706 specification standardizes the energy-efficiency measurement of various wireless network technologies, such as GSM/Edge, LTE, and WiMAX, considering both throughput and coverage area relative to energy consumption.

Although many standardization efforts exist, none apply directly to networks operating close to the fundamental limits. Future work could focus on defining a metric similar to the absolute energy efficiency metric to measure the efficiency of networks and devices relative to fundamental energy dissipation limits. Researchers devoted to multiterminal scenarios should develop a new absolute energy efficiency metric.

Regardless of network size, any standardization effort should consider the acceptable amount of CO_2 emissions. These standardization efforts, together with increased consumer awareness, will surely force manufacturers to build communication and computing devices with minimum energy dissipation and CO_2 emissions.

As IT solutions multiply, handling the adverse environmental effects of ICT systems is becoming more cumbersome. Although many approaches aim to reduce the energy consumption of ICT systems, a fundamental approach to achieving systems with minimum energy consumption is missing. We have provided but a glimpse of existing green communication efforts and studies of fundamental energy consumption per reliable bit. New research directions will inevitably lead to networks that operate at the fundamental energy consumption limit.

To develop environmentally friendly ICT systems, the first step is to determine the fundamental limits of CO_2 emissions for point-to-point communications. The next step is to design novel technologies to develop a physical layer operating at the fundamental energy and CO_2 limits. Then the immediate challenge is how to use a layered network architecture that is compatible with this physical layer.

Even after solving these immediate problems, the ICT community faces many obstacles in creating networks with minimal emissions, such as selecting the optimum routes minimizing CO_2 emissions or adaptively controlling the rate to minimize the use of emission-intensive network paths. Regardless of the work ahead, the dream of always-on communications in concert with a greener environment will continue to motivate future efforts. C

Acknowledgments

This work was supported in part by Turk Telekom under grant 11315-04 and by the IBM Shared University Research Award.

References

1. X.W. Dong, T. El-Gorashi, and J.M.H. Elmirghani, "Green IP over WDM Networks with Data Centers," *IEEE/OSA J. Lightwave Technology*, vol. 29, no. 12, pp. 1861-1880.
2. M. Parker and S. Walker, "Roadmapping ICT: An Absolute Energy Efficiency Metric," *IEEE/OSA J. Optical Comm. Networks*, Aug. 2011, pp. A49-A58.
3. R. Landauer, "Energy Requirements in Communication," *Applied Phys. Letters*, Dec. 1987, pp. 2056-2058.
4. A. Bérut et al., "Experimental Verification of Landauer's Principle Linking Information and Thermodynamics," *Nature*, Mar. 2012, pp. 187-189.
5. B. Sovacool, "Valuing the Greenhouse Gas Emissions from Nuclear Power: A Critical Survey," *Energy Policy*, vol. 36, no. 8, pp. 2940-2953.
6. A. Jain, S.R. Kulkarni, and S. Verdu, "Minimum Energy per Bit for Gaussian Broadcast Channels with Common Message and Cooperating Receivers," *Proc. 47th Ann. Allerton Conf. Communication, Control, and Computing*, IEEE, 2009, pp. 740-747.
7. S. Verdu, "On Channel Capacity per Unit Cost," *IEEE Trans. Information Theory*, Sept. 1990, pp. 1019-1030.
8. A.-C. Orgerie, L. Lefevre, and I. Guerin-Lassous, "Energy-Efficient Bandwidth Reservation for Bulk Data Transfers in Dedicated Wired Networks," *J. Supercomputing* (special issue on green networks), Mar. 2011, pp. 1-28.
9. D. Kliazovich, P. Bouvry, and S. Khan, "GreenCloud: A Packet-Level Simulator of Energy-Aware Cloud Computing Data Centers," *J. Supercomputing*, Nov. 2010; http://dx.doi.org/10.1007/s11227-010-0504-1.
10. M.D.S. Bhawan and J.L.N. Marg, "Recommendations on Approach Towards Green Telecommunications," 2011; http://trai.gov.in/WriteReadData/recommendation/documents/green_telecom-12.04.2011.pdf.
11. ETSI, "TS 102 706: Environmental Engineering (EE) Measurement Method for Energy Efficiency of Wireless Access Network Equipment," 2011; www.etsi.org/deliver/etsi_ts/102700_102799/102706/01.02.01_60/ts_102706v010201p.pdf.

Murat Kocaoglu *is pursuing an MSEE at Koc University's Next-Generation and Wireless Communications Laboratory (NWCL). His research interests include nanoscale communications, energy-efficient channel and network coding, and information theory. Kocaoglu received a BSEE with a minor in physics from Middle East Technical University, Ankara. He is an IEEE student member. Contact him at mkocaoglu@ku.edu.tr.*

Derya Malak *is a research assistant in Koc University's NWCL, where she is pursuing an MSEE. Her research interests include nanoscale communications, intrabody molecular communication networks, and network information theory. Malak received a BSEE from Middle East Technical University. She is an IEEE student member. Contact her at dmalak@ku.edu.tr.*

Ozgur B. Akan *is a professor in the Department of Electrical and Electronics Engineering at Koc University and director of the NWCL. His research interests include wireless communications, bioinspired communications, nanoscale and molecular communications, and information theory. He is an IEEE Senior Member and member of ACM. Contact him at akan@ku.edu.tr.*