Performance Optimization With Energy Packets

Erol Gelenbe, *Fellow, IEEE*, and Yunxiao Zhang, *Student Member, IEEE*

***Abstract*—We investigate how the flow of energy and the flow of jobs in a service system can be used to minimize the average response time to jobs that arrive according to random arrival processes at the servers. An interconnected system of workstations (WSs) and energy storage (ES) units that are fed with randomly arriving harvested energy is analyzed by means of the energy packet network (EPN) model. The system state is discretized and uses discrete units to represent the backlog of jobs at the WSs and the amount of energy that is available at the ES units. An energy packet (EP), which is the unit of energy, can be used to process one or more jobs at a WS, and an EP can also be expended to move a job from one WS to another one. The system is modeled as a probabilistic network that has a product-form solution for the equilibrium probability distribution of system state. The EPN model is used to solve two problems related to using the flow of energy and jobs in a multiserver system, so as to minimize the average response time experienced by the jobs that arrive at the system.**

***Index Terms*—Energy harvesting, energy packet network (EPN), G-networks, optimization, renewable energy.**

I. Introduction

LARGE numbers of heterogeneous digital devices and computer servers are being incorporated using the Internet of Things [1]–[5] to manage cities and various service activities [6], including environmental monitoring, health, security, vehicles, emergency evacuation, and smart grids [7]–[9]. Such systems must operate autonomously over long time spans and can benefit from energy harvesting from renewable energy sources, such as wind, liquid flows, photovoltaic, and ambient electromagnetic fields. In addition, such systems need energy storage (ES) to be able to smooth the effects over time of the intermittent sources of renewable energy [10]–[12]. Thus, there has been considerable interest in understanding how harvested energy can be used to optimize the consumption of energy and quality of service (QoS) of communication systems [13]–[15]. A framework of energy cooperation sharing in communication networks with energy harvesting was discussed in [16], while energy harvesting in a two-user cooperative Gaussian multiple access channel was considered in [17]. Some of the work until 2015 was reviewed in [18], regarding energy harvesting wireless communications and energy transfer from the perspective of communication and information theory. A queuing model of an energy efficient base station was presented in [19].

Sharing of power from a common rechargeable battery for different wireless channels was considered in [20]. Since the sustainability of information and computer technology improves with energy saving techniques [21], [22], much work was devoted to communication networks that manage energy consumption while meeting or optimizing QoS [23], [24]. Optimal routing policies for energy savings [25], [26], and optimal scheduling of data transmission for energy usage optimization [27] have also been studied. Energy efficient cloud servers and data centers are also very important [28], [29].

Motivated by these considerations, recent work has developed the energy packet network (EPN) paradigm [30]–[32], which is a discrete state-space modeling framework based on G-networks [33], which have a broad range of applications [34], [35] and can be used for evaluating both performance and energy consumption in a system where computer jobs, data in the form of packets, and energy represented by energy packets (EPs), interact in a complex, interconnected computer-communication system. This approach uses queuing theory, so that the joint behavior of discretized energy flows and the flows of computer jobs and data are analyzed within a single model. It was recently used for the analysis of the backhaul of mobile networks operating with intermittent renewable energy [36]. In previous work [37]–[39] optimization algorithms were developed on the basis of queuing networks, to dispatch network packets and minimize composite cost functions combining overall network energy consumption and QoS. The use of a central energy store (ES) was compared with a distributed storage facility in [40] with regard to overall efficiency, while a utility function, which combines throughput and the probability that the system does not run out of energy, was used in [41] for system optimization. The EPN model has also recently generated further interest [32], [42], [43] to optimize sensor networks and computer systems that operate with harvested energy.

A new product-form solution (PFS) (distinct from G-networks) was derived in [44] for a tandem network of N nodes using harvested energy stored in batteries; this analytical approach was initiated in [45] for single-node systems and developed in [46] for two-node systems. In addition, the work in [44] only applies to tandem networks (while in this paper we consider more general network structures) and furthermore [44] assumes that one EP can only serve to process exactly one job, while this paper discusses the case where one EP processes a batch of jobs.

Manuscript received June 4, 2018; revised October 28, 2018 and January 13, 2019; accepted February 12, 2019. Date of publication May 16, 2019; date of current version November 22, 2019. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme Project SDK4ED, under Grant 780572. *(Corresponding author: Erol Gelenbe.)*

E. Gelenbe is with the Institute of Theoretical and Applied Informatics of the Polish Academy of Sciences (IITIS-PAN), Gliwice 44-100, Poland, and also with the Intelligent Systems and Networks Group, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: e.gelenbe@imperial.ac.uk).

Y. Zhang is with the Intelligent Systems and Networks Group, Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: yunxiao.zhang15@imperial.ac.uk).

Digital Object Identifier 10.1109/JSYST.2019.2912013

Another work (also unrelated to G-networks) has proposed a practical hardware based design for switching and forwarding power and data simultaneously in a "power packet" system, which can be implemented on indoor power lines as well as on computer boards and chips [47], [48].

Here, we consider servers or workstations (WSs) that are powered by a battery or an ES, which is charged from a source of intermittent energy such as wind or photovoltaic. Energy leakage can also occur from an ES. Energy is represented by discretized EPs, and one EP is the smallest amount of energy represented in the system.

Thus, an EP is a basic unit of energy (for instance 100 W-s or 100 J), which is common to the system as a whole. With one EP we assume that *a WS can execute one or more jobs*. Thus, if a WS is more energy efficient, it will execute more jobs with a single EP. A WS (i.e., computer) that is more energy efficient will execute more jobs on average with a single EP. These assumptions generalize earlier work [41] where an EP was used to process a single job. On the other hand, this paper does not address the synchronization or dependencies between jobs in different WSs, as would occur when multiple jobs on different servers may be updating a shared set of data [49].

Specifically, we address the following two relevant problems of practical interest.

1) In Problem 1, we assume that EPs cannot be moved from one ES to another. Similarly, we assume that jobs cannot be moved from one WS to another. The system as a whole receives a total fixed power rate, expressed in EPs per second. Each single ES i is assigned to feed energy to a specific WS i where $i = 1, \ldots, N$; however, energy leakage can also occur from each ES. We are given the probability distribution of the number of jobs that a single EP can process at each given WS, and this distribution may differ at different WSs. The problem we solve is to select the fraction of power that is sent to each of the ESs so as to minimize the overall average response time R of the jobs in the system. If we denote by w_i (in 1/s) the maximum rate at which ES i feeds energy to WS i, then the peak power consumption of WS i is obviously also w_i EPs/s.
2) In Problem 2, again we have N ESs, each of which is allocated to its corresponding WS. We assume that EPs are allocated at a fixed rate to each ES. With probability D_i, we move a job that is at the head of the WS queue at node i, and if the job is moved it enters the queue at WS j with probability M_{ij}. The corresponding probability matrix is $\mathbf{M} = [M_{ij}]$. If a job is moved, then just moving it will consume one EP. Thus, the second problem considered is to select vector $D = (D_1, \ldots, D_N)$ so as to minimize R. Basically, this means that we are deciding whether to move a job or not from any of the WS i so as to reduce the workload at WS i and increase it at WS j, knowing that moving it will itself consume one EP at WS i, which then cannot be used to process another job. On the other hand, with probability $(1 - D_i)$ the decision will be to execute jobs locally rather than to move a job, in which case (as in Problem 1) a batch of jobs will be executed at WS i.

To solve these problems, we use the EPN model with time independent or stationary parameters, and we solve it in *steady state*.

Because the energy sources are time varying, one can ask whether a stationary model is useful. In fact, the time variations in energy harvesting, for photovoltaic or wind, would be in the tens of minutes, half-hour to hour (time of day) range. On the other hand, our model deals with millisecond up to tens of seconds time constants, which concern the execution times of computer programs. Thus, during the execution of hundreds to thousands of consecutive computer programs, the energy flow parameters will not change significantly, which is why we are justified in using a stationary model and in computing steady-state values. Therefore, over the longer time range, the optimizations described in this paper can be applied for different time-of-day effects, and the optimal parameters would be recomputed each time the energy flow parameters change.

Since in this paper one EP can be used to execute one or more jobs, the size of an EP has been chosen to be quite large. We could have also selected a "dual model" where one job is executed with one or more EPs, which would have been justified if an EP is a small unit of energy. The EPN paradigm admits both approaches, and in both the cases we can exploit the theory of G-Networks. However, in this paper we have just taken one of these two approaches, i.e., one EP is used to execute one or more jobs.

In the sequel, Section II summarizes some of the properties of G-Networks and shows how the EPN model is based on such models. The EPN model parameters are detailed in Section III. In Sections IV and V, we solve two optimization problems related to the allocation of jobs to different WSs based on their energy efficiency and the availability of energy, and we provide illustrative examples. Conclusions are drawn in Section VI.

II. EPN and Its G-Network Representation

The EPN system considered is schematically presented in Fig. 1. Jobs that must be executed in the system are modeled as ordinary customers in a queuing network. They arrive at one of the N WSs, say WS i, at a rate of λ_i jobs/s. Each WS is represented as a queue containing job. Jobs first arrive at a given WS i; each WS i has an ES battery denoted by ES i, and there are a total of N ESs. EPs arrive from an external intermittent energy source at rate γ_i EPs/s to ES i, which can be viewed as a "queue of EPs."

As shown in Fig. 1, in the EPN model, the EPs in ES i either can be forwarded to the corresponding WS i on demand with probability d_i, or moved to another ES node j with probability P_{ij} to balance the energy distribution. However, in the sequel, we assume that $d_i = 1$. The jobs in WS i can be processed locally with probability D_i or forwarded to some other WS j with probability M_{ij} for further steps of execution. In this figure, w_i is the rate at which EPs are forwarded from ES i to one of the WSs. On the other hand, δ_i is the loss rate of energy (i.e., leakage) from ES i.

We denote the number of jobs at WS i by $K_i(t)$, while $L_i(t)$ denotes the number of EPs at ES i, at time t. We assume that

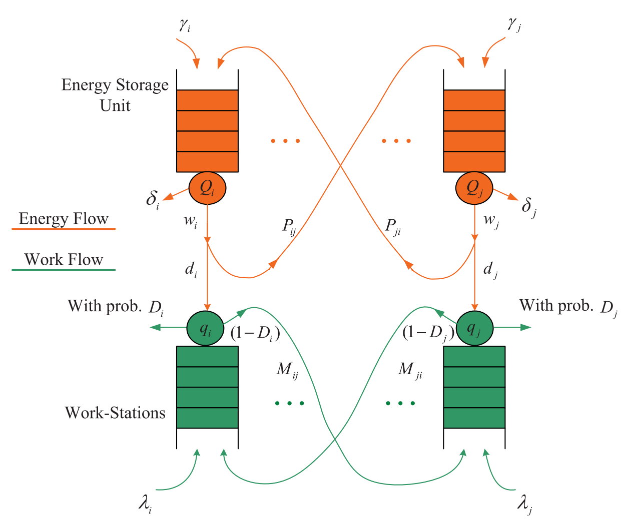


Fig. 1. Schematic representation of an EPN system with N WS (WS) nodes and N ES nodes. The EPs are accumulated in the ESs (amber), and jobs are accumulated in the WSs (green). The EPs in ES i either can be forwarded to the corresponding WS i on demand with probability d_i, or moved to another ES node j with probability P_{ij} to balance the energy distribution. The jobs in WS i can be processed locally with probability D_i or forwarded to some other WS j with probability M_{ij} for further steps of execution. In this figure, w_i is the rate at which EPs are forwarded from ES i to one of the WSs. On the other hand, δ_i is the loss rate of energy (i.e., leakage) from ES i.

both the WS queues, and the ES queues (i.e., batteries) are unbounded, i.e., of infinite capacity. EPs at ES i are expended because of energy leakage, consumed by the WSs, or moved in the following manner.

1) If $L_i(t) > 0$, then ES i will:
   a) either leak energy at some rate $\delta_i \geq 0$ EPs/s, and after a time of average value δ_i^{-1}, we will have one less EP at ES i because of energy leakage. The successive EP leakage times for the ith ES are modeled as independent and identically distributed (i.i.d.) random variables having a common exponential distribution with parameter δ_i;
   b) or ES i will forward one EP at rate w_i to WS i. A more general scheme is described in Fig. 1 where EPs are allowed to move between ESs.
2) Each EP is used locally by WS i as follows:
   a) with probability $1 \geq D_i \geq 0$, one EP will be expended to serve a batch of up to B_i jobs at WS i. If $K_i(t) > 0$, then the EP will serve $\min[K_i(t), B_i]$ jobs in one step and after the service we end up with $K_i(t^+) = K_i(t) - \min[K_i(t), B_i]$. Since each job may have different energy requirements at WS i, we assume that the number of jobs that can be processed with a single EP at WS i is a random variable;
   b) since our purpose is to model different WSs that have different levels of energy efficiency, a *single EP* is used to process one or more jobs, if there are jobs waiting in the WS queue;
   c) with probability $1 - D_i$, if $K_i(t) > 0$, one EP will be used to serve just one job, and then forward that job to another WS j according to the transition probability matrix $\mathbf{M} = [M_{ij}]$. As a result, we will have $K_i(t^+) = K_i(t) - 1$, $K_j(t^+) = K_j(t) + 1$;
   d) if an EP arrives at a WS i and $K_i(t) = 0$, then the EP will just be expended to keep the WS in working order (i.e., to keep it on), and no jobs will be processed or moved.

Thus, if $d_i = 1$ and $D_i = 1$, then the EPs at each ES i are only used locally to process the jobs at WS i, and keep WS i "ON" when there are no jobs to process.

A. G-Network Model

The EPN model discussed above is a special case of a family of queuing networks known as G-Networks, which were developed starting around 1990 [50], continuously over the years [51] including models for system security [52], to date [53]–[55]. A remarkable and useful property of a G-Network is the "PFS," which we recall at the end of this section and use to analyze the EPN.

The queuing model we discuss here corresponds to a multiclass G-Network with batch removal and multiple classes of customers [56], [57]. It is an open network containing a finite number v of queues or service stations, in which customers circulate. These customers can belong to one of C classes, so that each customer class can have different arrival rates to the network and can also have different routing probabilities within the network. Each of the C classes can contain customers of three types. These types are the "positive." "negative" and "triggers." Other types of customers that were developed more recently, e.g., "resets" [58] and "adders," are not used in this paper.

Positive customers are the normal queuing network customers, which request and obtain service at the queues. They belong to one of the C classes. We denote by $\kappa_{c,i}(t)$ the number of positive customers of class c at node i at time t. The *total number of positive customers* at node i at time t is denoted $K_i(t) = \sum_{c=1}^{C} \kappa_{c,i}(t)$.

At all of the v queues, positive customers have i.i.d. exponential service times of rates $r(1), \ldots, r(v)$, which are assumed in this paper to be identical for all classes of customers. After completing service and leaving a node i, a positive customer of class c can become as follows:

1) a positive customer of class c' at node j with probability $\Pi^+_{c,i,c',j}$, and we denote the corresponding transition probability matrix as $\mathbf{\Pi}^+ = [\Pi^+_{c,i,c',j}]$; or
2) the positive customer can leave the network with probability $l_{c,i}$; or
3) it can change into a negative customer of class c' and join node j with probability $\Pi^-_{c,i,c',j}$, in which case it will remove, or "instantaneously serve," a batch of positive customers of class c', and the batch is of maximum size $B_{c',j}$ at queue j. For the purpose of this paper, *we assume that the probability distribution of batch size $B_{c',j}$ does not depend on class c'*, so that $B_{c',j}$ is a random variable with the following probability distribution:

$$\pi_j(s) = \Pr[B_{c',j} = s] \geq 0, \ s \geq 1. \quad (1)$$

Thus, if the negative customer of class c at node i then arrives to queue j as a class c' customer at time t, then a total of max $[\kappa_{c',j}(t), B_{c',j}]$ positive customers of class c' will be instantaneously removed from the queue at j so that $\kappa_{c',j}(t^+) = 0$ if $B_{c',j} \geq \kappa_{c',j}(t)$, and $\kappa_{c',j}(t^+) = \kappa_{c',j} - B_{c',j}$ if $B_{c',j} < \kappa_{c',j}(t)$. Furthermore, the negative customer disappears at time t^+ after it has had its effect on the queue. Also, if $\kappa_{c',j}(t) = 0$, then the negative customer itself disappears, and no customer is removed from queue j;

4) finally, the positive customer of class c leaving queue i can become a "trigger" of class c' at queue j with probability $\Pi^T_{c,i,c',j}$, in which case it will move a class c' customer from queue j to queue l, and that customer becomes a class c'' customer at queue l, with probability $Q_{c',j,c'',l} \geq 0$. If queue j does not contain a class c' customer when the trigger arrives to queue j, then no customer is transferred from j to l, and the trigger disappears;
5) the effect of a negative customer and of a trigger are instantaneous: They occur in zero time; i.e., a negative customer or trigger arriving to a queue at time t will modify the queue's state at time t^+. Furthermore, both a negative customer and a trigger will themselves disappear after they have visited a queue;
6) queues also have *external* positive-, negative-, and trigger-type customer arrivals at rates $\lambda^+_{c,i}$, $\lambda^-_{c,i}$, and $\lambda^T_{c,i}$, which can differ for each class c and queue i, according to independent Poisson processes at each of the queues. Furthermore, externally arriving customers will have exactly the same effect at a queue as the ones that arrive from another queue;
7) positive customers at WS i have service times, which are mutually independent and exponentially distributed with rate $r(i)$; note that the service rate is same for any class c.

For all (c, i), the probabilities introduced above will satisfy the following:

$$l_{c,i} + \sum_{c'=1}^{C} \sum_{j=1}^{v} \left[\Pi^+_{c,i,c',j} + \Pi^-_{c,i,c',j} + \Pi^T_{c,i,c',j}\right] = 1 \tag{2}$$

$$\sum_{c''=1}^{C} \sum_{l=1}^{v} Q_{c,i,c'',l} = 1. \tag{3}$$

Let $\Lambda^+_{c,i}$, $\Lambda^-_{c,i}$, and $\Lambda^T_{c,i}$ denote the total arrival rate to queue i of class c customers that are of positive-, negative-, and trigger-type, respectively. Then, the "traffic equations" for the system are given by the following:

$$\begin{aligned}
\Lambda^+_{c,i} = \lambda^+_{c,i} &+ \sum_{c'=1}^{C} \sum_{j=1}^{v} r(j) q_{c',j} \Pi^+_{c',j,c,i} \\
&+ \sum_{c'=1}^{C} \sum_{j=1}^{v} r(j) q_{c',j} \sum_{c''=1}^{C} \sum_{l=1}^{v} \Pi^T_{c',j,c'',l} q_{c'',l} Q_{c'',l,c,i} \\
\Lambda^-_{c,i} = \lambda^-_{c,i} &+ \sum_{c'=1}^{C} \sum_{j=1}^{v} r(j) q_{c',j} \Pi^-_{c',j,c,i} \\
\Lambda^T_{c,i} = \lambda^T_{c,i} &+ \sum_{c'=1}^{C} \sum_{j=1}^{v} r(j) q_{c',j} \Pi^T_{c',j,c,i}
\end{aligned} \tag{4}$$

where

$$q_{c,i} = \frac{\Lambda^+_{c,i}}{r(i) + \Lambda^T_{c,i} + \Lambda^-_{c,i} \cdot \left[\frac{1 - \sum_{s=1}^{\infty} q^s_{c,i} \pi_i(s)}{1 - q_{c,i}}\right]}.$$

In the sequel, we will assume the following.

1) At any queue i only positive customers, negative customers, and triggers of *a specific single class c_i can arrive*.
2) Therefore, for a specific c_i we have: $\Lambda^T_{c,i} = \Lambda^-_{c,i} = \Lambda^+_{c,i} = 0$ if $c \neq c_i$.
3) Also, $\Lambda^T_{c_i,i} \geq 0$, $\Lambda^-_{c_i,i} \geq 0$, $\Lambda^+_{c_i,i} \geq 0$.

As a consequence we have

$$q_{c_i,i} = \frac{\Lambda^+_{c_i,i}}{r(i) + \Lambda^T_{c_i,i} + \Lambda^-_{c_i,i} \cdot \left[\frac{1 - \sum_{s=1}^{\infty} q^s_{c_i,i} \pi_i(s)}{1 - q_{c_i,i}}\right]}. \tag{5}$$

With these assumptions, the following result follows from the previously conducted work in [56] and [57]:

Result 1—PFS: Let $\mathbf{K}(t) = (K_1(t), \ldots, K_v(t))$. If the traffic equations in (4) have a unique solution such that all the $q_{c,i}$ in (5) lie between 0 and 1, i.e., $0 < q_{c,i} < 1$ for $1 \leq i \leq v$ and $1 \leq c \leq C$, then denoting by

$$q^*_i = \sum_{c_i=1}^{v} q_{c_i,i} \tag{6}$$

the following result holds:

$$\lim_{t \to \infty} \Pr\left[\mathbf{K}(t) = (k_1, \ldots, k_v)\right] = \prod_{i=1}^{v} [q^*_i]^{k_i} (1 - q^*_i). \tag{7}$$

Directly following from the above PFS (7), we can see that the marginal queue length probability distribution for any queue j is given by

$$\begin{aligned}
\lim_{t \to \infty} \Pr[K_j(t) = k_j] &= \sum_{i=1, i \neq j}^{v} \sum_{k_i=1,\, i \neq j}^{\infty} \left[\prod_{i=1}^{v} [q^*_i]^{k_i} (1 - q^*_i)\right] \\
&= [q^*_j]^{k_j} (1 - q^*_j).
\end{aligned} \tag{8}$$

III. EPN as a G-Network and Its Optimization

We now refer to the EPN of Fig. 1 and to the discussion in Sections II and II-A. The EPN in Fig. 1 can be represented by a G-Network with $v = 2N$ queues, where the WSs are represented by queues $1, \ldots, N$, while the ESs are represented by queues $N + 1, \ldots, 2N$.

Specifically, with regard to the notation in Section II and Section II-A, we have

1) The network has $C = 2$, i.e., two classes of customers: Class 1 refers to the jobs to be executed in the WSs. Class 2 refers to the EPs.
2) Note that negative customers and triggers cannot arrive at any of the queues from the outside world, i.e., $\lambda^-_{c,i} = \lambda^T_{c,i} = 0$ for $c = 1,\ 2$ and $i \in \{1, \ldots, 2N\}$.
3) Class 1 customers can only be "positive customers" and represent the jobs being served at the WSs. Hence, $\lambda^+_{1,i} = \lambda_i$ and $\lambda^+_{2,i} = 0$ for $i = 1, \ldots, N$.

4) Furthermore, jobs at the WSs are only removed, or moved to another WS, under the effect of EPs, i.e., $r(i) = 0$ and $l_{1,i} = l_{2,i} = 0$ for $i = 1, \ldots, N$.
5) EPs are positive customers at the ES queues $N + 1, \ldots, 2N$. Hence, for $i, j \in \{N+1, \ldots 2N\}$: $\lambda^+_{2,i} = \gamma_i$, $\lambda^-_{2,i} = 0$, $\lambda^+_{1,i} = \lambda^-_{1,i} = 0$, and $r(i) = w_i + \delta_i$.
6) Also, $\Pi^+_{2,i,2,j} = P_{ij}$ where P_{ij} is the probability that EPs are moved from ES i to ES j. However, in the system that we will analyze, EPs cannot be moved from one ES to another, so that $\Pi^+_{2,i,2,j} = 0$.
7) $\Pi^-_{2,i,2,j} = 0$ because EPs cannot "eliminate or destroy" other EPs.
8) $\Pi^-_{1,i,1,j} = \Pi^-_{1,i,2,j} = 0$ because jobs cannot eliminate other jobs or EPs.
9) Note that $l_{2,i} = \frac{\delta_i}{\delta_i + w_i}$ is the probability that an EP is leaked out of ES i rather than being forwarded to WS i.
10) EPs that leave ES (queue) $j = N + i$ and arrives at WS i, with $i \in \{1, \ldots, N\}$, becomes either negative customers (serving a batch of jobs) or triggers (moving a job to another queue), with probability $d_j . \frac{w_j}{\delta_j + w_j}$.
11) On arrival at WS i, $1 \le i \le N$, with probability D_i, an EP becomes a *negative customer with batch removal*, representing that an EP is used to process one or more jobs at the WS. The probability distribution of the size of the batch of jobs that can be served or "removed" is $\pi_i(s) = \Pr[B_i = s]$, and $\Pi^-_{2,j,1,i} = D_i . d_j . \frac{w_j}{\delta_j + w_j}$, with $j \in \{N+1, \ldots, 2N\}$ and $i = j - N$.
12) On arrival at i, with probability $1 - D_i$, an EP becomes a *trigger*, so that $\Pi^T_{2,j,1,i} = (1 - D_i) d_j . \frac{w_j}{\delta_j + w_j}$, and $Q_{1,i,1,m} = M_{im}$, for $j \in \{N+1, \ldots, 2N\}$, $i = j - N$, $1 \le m \le N$.
13) Note that $\Pi^T_{2,j,2,i} = \Pi^T_{1,j,2,i} = \Pi^T_{1,j,1,i} = 0$ for all $i, j \in \{1, \ldots, 2N\}$, and $\Pi^T_{2,j,1,i} = 0$ if $i \ne j - N$ for $N + 1 \le j \le 2N$.
14) $\Pi^+_{1,i,1,j} = (1 - D_i) M_{ij}$, $\Pi^+_{1,i,2,j} = 0$, $\Pi^+_{2,i,1,j} = 0$, $l_{1,i} = 0$, for $i, j \in \{1, \ldots, N\}$.
15) $l_{1,i} = 0$, $l_{2i} = 0$ for $i = 1, \ldots, N$, and $l_{1i} = 0$, $l_{2,i} = \frac{\delta_i}{\delta_i + w_i}$ for $i = N + 1, \ldots, 2N$.
16) $1 - d_i = \sum_{j=1}^N P_{ij}$ for $i = 1, \ldots, N$, and $\sum_{j=1}^N M_{ij} = 1$ for $i = 1, \ldots, N$.

With regard to (5) of the G-Network model, the corresponding expressions for the EPN model are given for Classes 1 and 2 by the following expression:

$$q_{1,i} = \frac{\Lambda^+_{1,i}}{q_{2,i+N} w_i d_i \left[(1 - D_i) + D_i \frac{1 - \sum_{s=1}^{\infty} q^s_{1,i} \pi_i(s)}{1 - q_{1,i}} \right]} \quad (9)$$

where

$$\Lambda^+_{1,i} = \lambda_i + \sum_{j=1}^{N} q_{1,j} (1 - D_j) d_j w_j M_{ji} q_{2,j+N}$$

and

$$q_{2,i+N} = \frac{\gamma_i + \sum_{j=1}^{N} w_j q_{2,j+N} P_{ji}}{w_i + \delta_i}. \quad (10)$$

A. PFS for the EPN Model

Because the EPN model we have described is a special case of a G-Network with two classes of customers, namely jobs for Class 1, and EPs for Class 2, we can directly apply the PFS of (7). For this case, i.e., where we model an EPN, each of the queues is either a WS or an ES. WSs only contain Class 1 customers, and ESs only contain Class 2 customers.

Here, v of (7) has value $v = 2N$, and queues $1, \ldots, N$ are WSs, while the queues $N + 1, \ldots, 2N$ are the ESs.

As a consequence, the value q^*_i of (7) is given by

$$q^*_i = q_{1,i},\ 1 \le i \le N \quad q^*_i = q_{2,i},\ N + 1 \le i \le 2N. \quad (11)$$

Therefore

$$\lim_{t\to\infty} \Pr[\mathbf{K}(t) = (k_{1,1}, \ldots, k_{1,N}, k_{2,N+1}, \ldots, k_{2,2N})]$$
$$= \prod_{i=1}^{N} q^{k_{1,i}}_{1,i} (1 - q_{1,i}) q^{k_{2,i+N}}_{2,i+N} (1 - q_{2,i+N}). \quad (12)$$

If (9) and (10) have a unique solution such that all the $0 < q_{c,i} < 1$, for $1 \le i \le 2N$ and $1 \le c \le 2$. The marginal probability of the queue length for queue i and class $c = 1, 2$ is

$$\lim_{t\to\infty} \Pr[K_{c,i}(t) = k_{c,i}] = q^{k_{c,i}}_{c,i} (1 - q_{c,i}). \quad (13)$$

B. Cost Function, Parameters, and Optimization

Here, we will address two related optimization problems, which are outlined below. The objective is to minimize the average response time for jobs that come into the system, where the jobs arrive from the outside world to WS i at a given rate λ_i. Furthermore, the total arrival rate of EPs is fixed at some value γ, and each of the ESs has a transfer rate of EPs to the corresponding WS given by WS i and a local energy leakage rate δ_i, for $i = 1, \ldots, N$. Note that the transfer times for EPs from ESs to the corresponding WS are i.i.d. and exponentially distributed random variables with parameter WS i. Similarly, the successive leakage times for the EPs in the ith ES are also i.i.d. and exponentially distributed with parameter δ_i.

To simplify the analysis, we make an assumption regarding the probability distribution $\pi_i(s)$. Specifically, we assume that

$$\pi_i(s) = (1 - u_i) u_i^{s-1}, 0 < u_i < 1, s \ge 1, \sum_{s=1}^{\infty} (1 - u_i) u_i^{s-1} = 1.$$

The *average of the maximum number of jobs that can be processed by a single EP at WS* i is

$$E[B_{c_i,i}] = \sum_{s=1}^{\infty} s(1 - u_i) u_i^{s-1} = \frac{1}{1 - u_i}. \quad (14)$$

Although this geometric assumption regarding the probability of the number of jobs being serviced by a single EP is convenient for computational purposes, analytical results can also be obtained for general distributions when $q_{1,i}$ are quite large and, hence, close to one for a heavily loaded system, or very small and close to zero for a lightly loaded system.

1) Problem 1: Consider the case where the EPs cannot move between ESs so that $P_{ji} = 0$ and $d_i = 1$. Also, assume that jobs cannot be moved between WSs, i.e., $D_i = 1$. In this case, assume that the total renewable energy flow into WS i is $\gamma_i = p_i.\gamma$.

The cost function that needs to be minimized represents the overall average job response time as follows:

$$R = \frac{1}{\sum_{i=1}^{N} \lambda_i} \sum_{i=1}^{N} \frac{q_{1,i}}{1 - q_{1,i}}. \tag{15}$$

Regarding (9) and (10) with the specific restrictions for this case with $d_i = 1,\ D_i = 1$, for $1 \leq i \leq N$, we have

$$q_{1,i} = \frac{\lambda_i}{q_{2,i} w_i \left[\frac{1 - \sum_{s=1}^{\infty} q_{1,i}^s \cdot \pi_i(s)}{1 - q_{1,i}}\right]} \tag{16}$$

$$q_{2,i+N} = \frac{\gamma p_i}{w_i + \delta_i}. \tag{17}$$

Furthermore, there is only one class of customers (the computer jobs) at WSs, i.e., queues $1, \ldots, N$, and similarly just one class of customers (the EPs) at the ESs, i.e., queues $N+1, \ldots, 2N$; we can write: $q_i^* = q_{1,i}$ and $q_{i+N}^* = q_{2,i+N}$ for $i \in \{1, \ldots, N\}$.

Problem 1: is then to choose $p = (p_1, \ldots, p_N)$ so as to minimize R for a given value of γ and for given energy leakage rate δ_i at each ES i.

2) Problem 2: In the second problem, we assume that $d_i = 1,\ i = 1, \ldots, N$ so that EPs stay in the same ES unit where they have been initially allocated. However, in Problem 2, we do allow jobs to be moved between WSs, and their movement is specified via a fixed probability matrix $\mathbf{M} = [M_{ij}]$, where M_{ij} is the probability that a job that is currently at WS i is moved for execution to WS j.

Recall that D_i is the probability that at station i the job at the head of the queue is allowed to move to station j with probability M_{ij}. Note that in this case, because the jobs do move, the average response time R will be based on the total effective arrival rate of jobs to each WS, including the jobs arriving from other WSs. Thus,

Problem 2: is to find the value of $D = (D_1, \ldots, D_N)$ that minimizes R, i.e., the overall average response time of jobs, for a given fixed movement matrix $\mathbf{M}$.

IV. Analysis of Problem 1

Using the Little's Formula we write

$$R = \frac{1}{\lambda^+} \sum_{i=1}^{N} \frac{q_i^*}{1 - q_i^*} \tag{18}$$

where $\lambda^+ = \sum_{i=1}^{N} \lambda_i$.

Note that $\Lambda_{1,i}^+ = \lambda_i$ when $D_i = 1$ for all $i = 1, \ldots, N$. Substituting $\frac{(1-u_i)u_i^s}{u_i}$ into (16), we have

$$\begin{aligned} q_i^* &= \frac{\lambda_i}{w_i q_{i+N}^*} \times \left[\frac{1 - \sum_{s=1}^{\infty} \frac{(1-u_i)u_i^s}{u_i} q_i^{*s}}{1 - q_i^*}\right]^{-1} \\ &= \frac{\lambda_i}{u_i \lambda_i + w_i q_{i+N}^*}. \end{aligned} \tag{19}$$

Substituting (19) into the cost function R, we obtain

$$R = \frac{1}{\lambda^+} \sum_{i=1}^{N} \frac{\lambda_i}{\sigma_i \gamma p_i + \lambda_i (u_i - 1)} \tag{20}$$

where

$$\sigma_i = \frac{w_i}{w_i + \delta_i} \tag{21}$$

denotes the energy efficiency with regard to the leakage of the ith ES node.

Choosing $p_i \geq 0$ so as to minimize R is an optimization problem subject to constraint $\sum_{i=i}^{N} p_i = 1$. Therefore, we use Lagrange multipliers with the following Lagrangian:

$$\mathcal{L} = R + \beta \left(\sum_{i=1}^{N} p_i - 1\right). \tag{22}$$

Here, the Lagrange multiplier β is a real number.

Suppose $\mathbf{p}^* = (p_1^*, \ldots, p_N^*)$ is a local solution of the optimization problem. Then, the necessary Kuhn–Tucker conditions are

$$\nabla_p \mathcal{L}(p^*, \beta^*) = 0 \tag{23}$$

and

$$\sum_{i=1}^{N} p_i^* - 1 = 0 \tag{24}$$

where $\mathbf{p}^*$ is a regular point for the constraint.

Solving for (23), we know that

$$\frac{\partial R}{\partial p_i} = \frac{-\lambda_i \sigma_i \gamma}{\lambda^+ \left[\sigma_i \gamma p_i + \lambda(u_i - 1)\right]^2} = -\beta \tag{25}$$

must hold. Then, rearranging (25), the solution p_i^* is

$$p_i^* = \frac{\lambda_i (1 - u_i)}{\sigma_i \gamma} + \sqrt{\frac{\lambda_i}{\lambda^+ \sigma_i \gamma \beta}}. \tag{26}$$

Moreover, the second necessary condition

$$\sum_{i=1}^{N} \left(\frac{\lambda_i (1 - u_i)}{\sigma_i \gamma} + \sqrt{\frac{\lambda_i}{\lambda^+ \sigma_i \gamma \beta}}\right) = 1 \tag{27}$$

also must hold. Solving (26) and (27) simultaneously, we obtain the following:

Result 2: The optimal solution to Problem 1 is given by

$$p_i^* = \frac{\lambda_i (1 - u_i)}{\sigma_i \gamma} + \frac{\sqrt{\frac{\lambda_i}{\sigma_i}}}{\sum_{i=1}^{N} \sqrt{\frac{\lambda_i}{\sigma_i}}} \left(1 - \sum_{i=1}^{N} \frac{\lambda_i (1 - u_i)}{\sigma_i \gamma}\right). \tag{28}$$

TABLE I
PARAMETERS IN PROBLEM 1

Parameters	Values
γ	150 EPs/sec
$\lambda_1,\ \lambda_2,\ \lambda_3$	$50, 30, 10$ jobs/sec
D_1, D_2, D_3	1, 1, 1
$w_1,\ w_2,\ w_3$	$100, 80, 50$ EPs/sec
$u_1,\ u_2,\ u_3$	$0.2, 0.2, 0.2$
M_{ij} for all i, j	0
P_{ij} for all i, j	0
$\delta_1,\ \delta_2,\ \delta_3$	$10, 8, 6$ EPs/sec
$d_1,\ d_2,\ d_3$	$1, 1, 1$

However, the sufficient condition that there exists an optimum solution p^* also needs to be examined. To guarantee the existence of the strict constrained local minimum, Hessian $\nabla_{pp}\mathcal{L}$ must be positive definite. Notice that $\nabla_{pp}\mathcal{L}$ is a diagonal matrix with diagonal entries

$$\frac{\partial^2 \mathcal{L}(p^*, \beta^*)}{\partial p_i^2} = \frac{\partial^2 R}{\partial p_i^2} = \frac{2\lambda_i \sigma_i^2 \gamma^2}{\lambda^+ \left[\sigma_i \gamma p_i^* + \lambda_i (u_i - 1)\right]^3}. \tag{29}$$

Thus, the sufficient condition holds if inequality

$$\sigma_i \gamma p_i^* > \lambda_i (1 - u_i) \tag{30}$$

is satisfied for all $i = 1, \ldots, N$. Substituting p_i^* into (30), we see that the inequality is equivalent to the following expression.

Result 3: The necessary condition for the optimal solution of Result 2 is given by

$$\gamma > \sum_{i=1}^{N} \frac{\lambda_i}{\sigma_i}(1 - u_i). \tag{31}$$

This condition is physically meaningful since it implies that the total rate of harvested EPs has to be sufficiently large so as to provide enough energy to power all the WSs, despite the energy leakage that also will occur at each ES.

Note from (14) that $1 - u_i = [\ E[B_{c_i,i}]\ ]^{-1}$, i.e., $(1 - u_i)$ is the inverse of the average of the maximum number of jobs that WS i can process with a single EP.

A. Example

In order to illustrate the analytically obtained optimal solution of Problem 1, we will consider a numerical example with three pairs of WS and ES nodes and the parameters shown in Table I.

We first examine the sufficient condition with respect to (30) to find the range of p_1, p_2, and p_3, respectively, and to guarantee that every ES can provide sufficient power to its corresponding WS. The numerical conditions are

$$0.2933 < p_1 < 1$$
$$0.1760 < p_2 < 1$$
$$0.0597 < p_3 < 1$$

with constraint $p_1 + p_2 + p_3 = 1$. Then, we calculate the values of delay R with all (p_1, p_2, p_3) and compare them to the optimal solution given in (28). The results are shown in Figs. 2 and 3 in which the *x*-axis and the *y*-axis are p_1 and p_2, while p_3 follows from $p_3 = 1 - p_1 - p_2$.

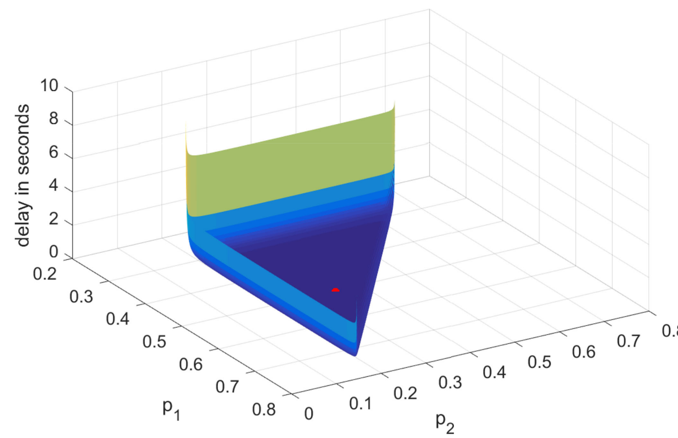


Fig. 2. Average job response time R for all (p_1, p_2) pairs. Note that range p_i for all i is not [0, 1] because of the constraints and the sufficient conditions.

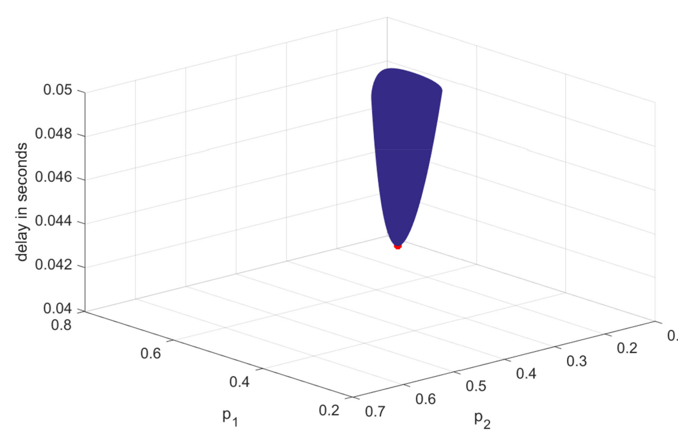


Fig. 3. Neighborhood of the optimum point at a much smaller scale of the average response time R along the *z*-axis.

Hence, a three-dimensional plot can be used to illustrate the relation of the average overall response time R and probability (p_1^*, p_2^*, p_3^*). The theoretical result from (28) gives the optimal solution $(p_1^*, p_2^*, p_3^*) = (0.5049, 0.3399, 0.1552)$, which produces the minimal overall delay $W = 42.9$ ms.

V. ANALYSIS OF PROBLEM 2

Here, as before, EPs from ES i are only consumed at WS i. Furthermore, from (9) and (10), we obtain the steady-state probabilities $q_{1,i}$ of the WS queues being nonempty, as well as the probabilities $q_{2,i}$ of the ESs being nonempty, as follows:

$$q_{1,i} = \frac{\Lambda_{1,i}^+ (1 - u_i q_{1,i})}{q_{2,i+N} w_i [1 - u_i q_{1,i} + u_i q_{1,i} D_i]} \tag{32}$$

$$q_{2,i+N} = \frac{\gamma_i}{w_i + \delta_i} \tag{33}$$

where

$$\Lambda_{1,i}^{+} = \lambda_i + \sum_{j=1}^{N} q_{1,j}(1-D_j)d_j w_j M_{ji} q_{2,j+N} \tag{34}$$

and we have used

$$\pi_{1,i}(s) = \frac{(1-u_i)u_i^s}{u_i},\ 0 < u_i < 1. \tag{35}$$

Note that by using the Little's Formula, we again have

$$R = \frac{1}{\sum_{i=1}^{N}\lambda_i}\sum_{i=1}^{N}\frac{q_{1,i}}{1-q_{1,i}} \tag{36}$$

but of course $q_{1,i}$ will be different. The following partial derivatives with respect to $D_k,\ k=1,\ldots,N$ will, therefore, be needed:

$$\begin{aligned}\frac{\partial q_{1,i}}{\partial D_k} = &\frac{1}{q_{2,i+N}w_i[1-u_i q_{1,i}+u_i q_{1,i}D_i]^2}\\ &\times\Bigg([1-u_i q_{1,i}+u_i q_{1,i}D_i]\Bigg[-u_i\Lambda_{1,i}^{+}\frac{\partial q_{1,i}}{\partial D_k}\\ &+(1-u_i q_{1,i})\sum_{j=1}^{N}\Bigg(q_{2,j+N}w_j M_{ji}(1-D_j)\frac{\partial q_{1,j}}{\partial D_k}\\ &-q_{2,j+N}w_j M_{ji}q_{1,j}\frac{dD_j}{dD_k}\Bigg)\Bigg]\\ &+\Lambda_{1,i}^{+}u_i(1-u_i q_{1,i})\left[(1+D_i)\frac{\partial q_{1,i}}{\partial D_k}+q_{1,i}\frac{dD_i}{dD_k}\right]\Bigg).\end{aligned} \tag{37}$$

Note that

$$\frac{dD_i}{dD_k} = \begin{cases}1, & \text{if } i=k\\ 0, & \text{otherwise.}\end{cases} \tag{38}$$

Rearranging (37), we have

$$\begin{aligned}A_i\frac{\partial q_{1,i}}{\partial D_k} = &\sum_{j=1}^{N} q_{2,j+N}w_j M_{ji}(1-D_j)\frac{\partial q_{1,j}}{\partial D_k}\\ &-q_{2,j+N}w_j M_{ji}q_{1,j}\frac{dD_j}{dD_k}+B_i\frac{dD_j}{dD_k}\end{aligned} \tag{39}$$

where

$$\begin{aligned}A_i &= \Lambda_{1,i}^{+}\left[\frac{1}{q_{1,i}}+\frac{u_i}{1-u_i q_{1,i}}-\frac{u_i(1+D_i)}{1-u_i q_{1,i}+u_i q_{1,i}D_i}\right]\\ B_i &= \frac{\Lambda_{1,i}^{+}u_i q_{1,i}}{1-u_i q_{1,i}+u_i q_{1,i}D_i}.\end{aligned}$$

Let us use the conventional notation $\mathrm{diag}(x_1,\ldots,x_N)$ for the diagonal matrix with diagonal entries of $(x_1,\ldots,x_N)$. Furthermore, we define the following matrices:

$$\begin{aligned}\mathbf{A} &= \mathrm{diag}(A_1,\ldots,A_N)\\ \mathbf{B} &= \mathrm{diag}(B_1,\ldots,B_N)\\ \mathbf{C_q} &= \mathrm{diag}(q_{2,1}w_1(1-D_1),\ldots,q_{2,N}w_N(1-D_N))\\ \mathbf{C_D} &= \mathrm{diag}(q_{2,1}w_1q_{1,1},\ldots,q_{2,N}w_Nq_{1,N}).\end{aligned}$$

Using the above notation, by augmenting scalars $\partial q_{1,i}/\partial D_k$ and dD_i/dD_k into a vector representation as follows:

$$\frac{\partial \mathbf{q_1}}{\partial D_k} \quad \frac{\partial \mathbf{D}}{\partial D_k} \tag{40}$$

we obtain

$$\frac{\partial \mathbf{q_1}}{\partial D_k} = [\mathbf{A}-\mathbf{M}^T\mathbf{C_q}]^{-1}[\mathbf{B}-\mathbf{M}^T\mathbf{C_D}]\frac{\partial \mathbf{D}}{\partial D_k}. \tag{41}$$

Moreover, define $\mathbf{J_R} = \nabla_{q_1}R$ as the gradient of R with respect to the elements of vector $\mathbf{q_1}$, or as follows:

$$\mathbf{J_R} = \left[\tfrac{1}{\lambda^{+}(1-q_{1,1})^2},\ldots,\tfrac{1}{\lambda^{+}(1-q_{1,N})^2}\right] \tag{42}$$

which is a $1\times N$ Jacobian matrix. By the chain rule, the gradient of the average response time, R, with respect to D_k is

$$\frac{\partial R}{\partial D_k} = \mathbf{J_R}[\mathbf{A}-\mathbf{M}^T\mathbf{C_q}]^{-1}[\mathbf{B}-\mathbf{M}^T\mathbf{C_d}]\frac{\partial \mathbf{D}}{\partial D_k}. \tag{43}$$

Since R is continuous and differentiable, gradient descent is useful for this optimization problem. At a given operation point $X_R = (\gamma, w, \delta, u, \lambda, M)$, the gradient descent algorithm at its mth computational step is

$$D_k^{(m+1)} = D_k^{(m)} - \alpha\frac{\partial R}{\partial D_k}\Big|_{D_k=D_k^{(m)}} \tag{44}$$

where $\alpha > 0$ is the rate of descent. The steps of the gradient algorithm are as follows:

1) initialize vector $\mathbf{D}$ and choose α;
2) solve the nonlinear equations given in (32) and (33) to yield steady-state utilizations $q_{1,i}$ and $q_{2,i}$;
3) calculate the partial derivatives as given by (43);
4) update the control parameter D_i using (44);
5) go to the Step 2 (above) until a sufficient number of iterations have been made, so that the difference between the absolute difference in the values of R in successive iterations is smaller than a preset value $\epsilon > 0$.

Note that in practice, this approach can also be used to apply *gradual* optimization of the system, since D_i are progressively modified, while the system may operate normally and slowly shifts toward the optimum.

A. Example

In order to illustrate the type of system that can be optimized, we consider *a remote sensing station*, which is powered by energy harvesting devices with three ES nodes, each of which powers a specific WS node, as shown in Fig. 4 in which ES i forwards EPs to WS i. For instance:

1) WS 1 is a server for the main sensor (e.g., a radar), with a local arrival rate λ_1 of jobs;

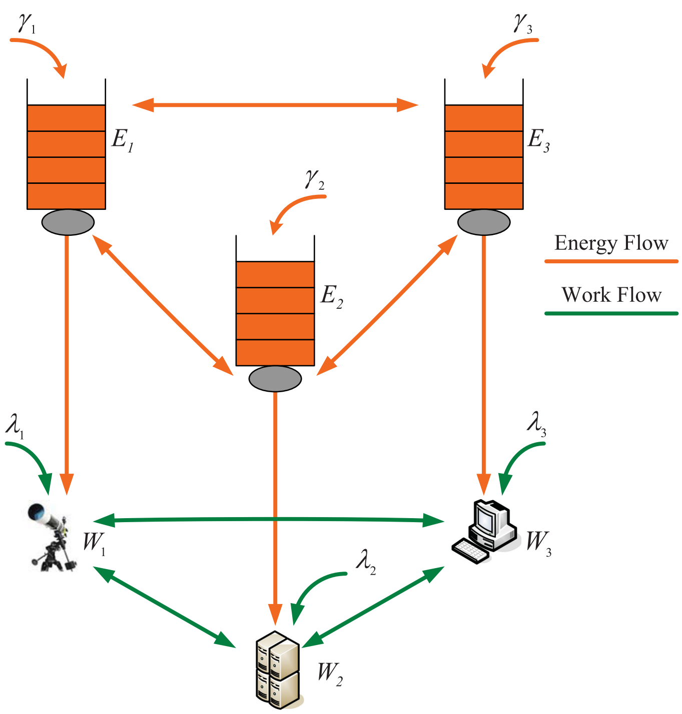


Fig. 4. Schematic representation of an EPN system with three WS and ES nodes, which models a remote sensing facility. WS 1 may represent the compute server for a radar or other sensor. WS 2 may be a communication server transmitting data to the external world, and WS 3 may be monitoring the temperature and security conditions at the remote station.

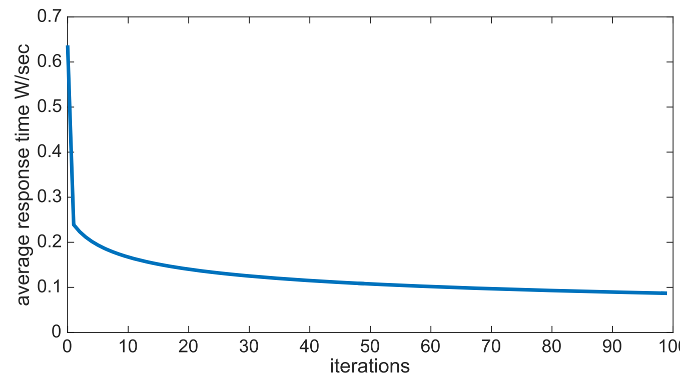


Fig. 5. Average response time R decreases and reaches its minimal value of 86.9 ms using the gradient descent algorithm.

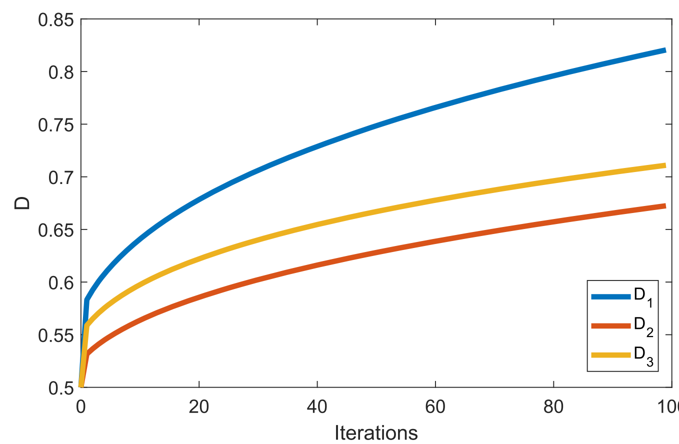


Fig. 6. Changes in the values of parameters D_1, D_2, and D_3 during the gradient descent.

TABLE II
PARAMETERS IN PROBLEM 2

Parameters	Values
$\gamma_1,\ \gamma_2,\ \gamma_3$	$50, 40, 40$ EPs/sec
$\lambda_1,\ \lambda_2,\ \lambda_3$	$30, 20, 10$ jobs/sec
P_{ij} for all i, j	0
$w_1,\ w_2,\ w_3$	$100, 80, 50$ EPs/sec
$u_1,\ u_2,\ u_3$	$0.3, 0.2, 0.1$
M_{ij} for all i, j	fixed
$\delta_1,\ \delta_2,\ \delta_3$	$10, 15, 20$ EPs/sec
$d_1,\ d_2,\ d_3$	$1, 1, 1$
α	0.01

2) WS 2 is a communication device, which is used to link the remote sensing location with the outside world, with an arrival rate of jobs λ_2;
3) WS 3 is a monitoring server, which processes sensor data for environmental data or security, with an arrival rate of jobs λ_3.

The EPs stay in the ES where they were initially stored. However, jobs are moved between the WSs to minimize the average time response time. Parameters are as shown in Table II, and the matrix M is chosen as follows:

$$M = \begin{bmatrix} 0.10 & 0.45 & 0.45 \\ 0.45 & 0.10 & 0.45 \\ 0.45 & 0.45 & 0.10 \end{bmatrix}. \tag{45}$$

When we apply the gradient descent algorithm with initial value $\mathbf{D} = (0.5, 0.5, 0.5)$, after 100 iterations we are close to the optimal value, $\mathbf{D}^* = (0.822, 0.673, 0.712)$, and R is reduced from 636.9 to 86.9 ms, as shown in Figs. 5, 6, and 7.

VI. CONCLUSION

In this paper, we have considered an EPN model representing a system where jobs can be moved between WSs, while EPs arrive at a WS from the ES directly associated with the WS. We have considered the case where the number of jobs serviced by a single EP is represented by a probability distribution. We also assume that each ES is subject to the loss of energy through leakage. Each WS may consume a different amount of energy per job that is processed, with respect to other WSs. We also assume that the WSs will consume energy even when they are idle.

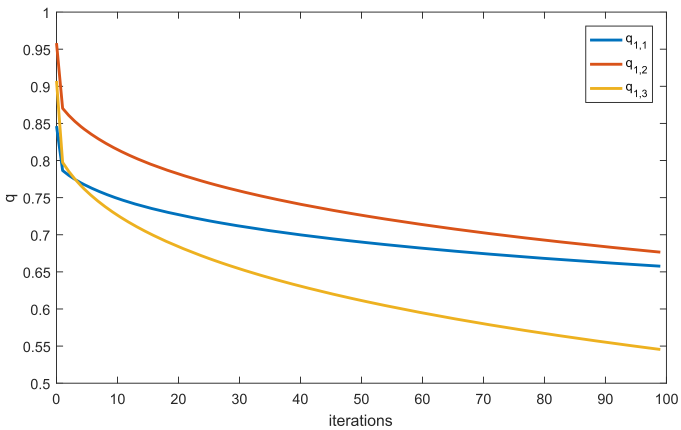


Fig. 7. Change in utilizations $q_{1,1}$, $q_{1,2}$, and $q_{1,3}$ during the gradient descent. The system remains stable since these values remain between 0 and 1.

We have first considered the case where neither jobs nor EPs can be moved, so that each WS executes locally the jobs that it receives, using energy from its own ES. In this case, we have considered how a common flow of EPs generated from a renewable energy source, should be distributed optimally among the ESs so that the average response time to the jobs can be minimized. This problem has been solved analytically for a special class of probability distributions for the number of jobs processed with one EP.

Then, with the same cost function to be minimized, we have considered the case where jobs can be moved among WSs, according to a given probability transition matrix, but each station can decide whether to move a job or not on the basis of the local decision probability D_i at WS i. In this case, again EPs that are allocated to a given ES are either consumed by the local WS or they are lost through leakage. Here, the optimization problem is to select the decision to move a job or not from a station where it is in queue to another station using the vector $\mathbf{D} = (D_1, \ldots, D_N)$. In this case, the solution is provided using a gradient descent algorithm of computational complexity $O(N^3)$. For both the problems, we have provided a numerical example to illustrate the results.

Future work will investigate the minimization of a cost function that combines the average response time of jobs, and the energy wastage through leakage or due to idle WSs that consume energy even when they do not process jobs.

REFERENCES

[1] C. Perera, C. H. Liu, S. Jayawardena, and M. Chen, "A survey on Internet of Things from industrial market perspective," *IEEE Access*, vol. 2, pp. 1660–1679, 2014.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.

[3] A. Whitmore, A. Agarwal, and L. Da Xu, "The Internet of Things: A survey of topics and trends," *Inf. Syst. Frontiers*, vol. 17, no. 2, pp. 261–274, 2015.

[4] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, 2010.

[5] L. Da Xu, W. He, and S. Li, "Internet of Things in industries: A survey," *IEEE Trans. Ind. Inform.*, vol. 10, no. 4, pp. 2233–2243, Nov. 2014.

[6] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[7] H. Bi and O. H. Abdelrahman, "Energy-aware navigation in large-scale evacuation using g-networks," *Probab. Eng. Informational Sci.*, vol. 32, pp. 340–352, 2018.

[8] E. Gelenbe and F.-J. Wu, "Future research on cyber-physical emergency management systems," *Future Internet*, vol. 5, no. 3, pp. 336–354, 2013.

[9] A. Al-Ali and R. Aburukba, "Role of Internet of Things in the smart grid technology," *J. Comput. Commun.*, vol. 3, no. 5, pp. 229–233, 2015.

[10] N. S. Wade, P. C. Taylor, P. D. Lang, and P. R. Jones, "Evaluating the benefits of an electrical energy storage system in a future smart grid," *Energy Policy*, vol. 38, no. 11, pp. 7180–7188, 2010.

[11] M. Black and G. Strbac, "Value of bulk energy storage for managing wind power fluctuations," *IEEE Trans. Energy Convers.*, vol. 22, no. 1, pp. 197–205, Mar. 2007.

[12] F. K. Shaikh, S. Zeadally, and E. Exposito, "Enabling technologies for green internet of things," *IEEE Syst. J.*, vol. 11, no. 2, pp. 983–994, Jun. 2017.

[13] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[14] A. Rahimi, Ö. Zorlu, A. Muhtaroglu, and H. Kulah, "Fully self-powered electromagnetic energy harvesting system with highly efficient dual rail output," *IEEE Sensors J.*, vol. 12, no. 6, pp. 2287–2298, Jun. 2012.

[15] E. Gelenbe, "A sensor node with energy harvesting," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 42, no. 2, pp. 37–39, 2014.

[16] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[17] B. Gurakan, O. Kaya, and S. Ulukus, "Energy and data cooperative multiple access channel with intermittent data arrivals," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2016–2028, Mar. 2018.

[18] S. Ulukus *et al.*, "Energy harvesting wireless communications: A review of recent advances," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[19] I. Dimitriou, S. Alouf, and A. Jean-Marie, "A Markovian queueing system for modeling a smart green base station," in *Proc. Eur. Workshop Perform. Eng*. Springer, 2015, pp. 3–18.

[20] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 581–593, Feb. 2010.

[21] W. van Heddeghem, S. Lambert, B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Trends in worldwide ICT electricity consumption from 2007 to 2012," *Comput. Commun.*, vol. 50, pp. 64–76, 2014.

[22] E. Gelenbe and Y. Caseau, "The impact of information technology on energy consumption and carbon emissions," *Ubiquity*, vol. 2015, pp. 1–15, 2015.

[23] C. C. Coskun, K. Davaslioglu, and E. Ayanoglu, "An energy-efficient resource allocation algorithm with QoS constraints for heterogeneous networks," in *Proc. Global Commun. Conf.*, Dec. 2015, pp. 1–7.

[24] O. Orhan, D. Gündüz, and E. Erkip, "Throughput maximization for an energy harvesting communication system with processing cost," in *Proc. IEEE Inf. Theory Workshop*, 2012, pp. 84–88.

[25] S. Sarkar, M. H. R. Khouzani, and K. Kar, "Optimal routing and scheduling in multihop wireless renewable energy networks," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1792–1798, Jul. 2013.

[26] Z. Mao, C. E. Koksal, and N. B. Shroff, "Near optimal power and rate control of multi-hop sensor networks with energy replenishment: Basic limitations with finite energy and data storage," *IEEE Trans. Autom. Control*, vol. 57, no. 4, pp. 815–829, Apr. 2012.

[27] M. A. Antepli, E. Uysal-Biyikoglu, and H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1712–1731, Sep. 2011.

[28] L. Newcombe, "Data centre energy efficiency metrics: Existing and proposed metrics to provide effective understanding and reporting of data centre energy," Brit. Comput. Soc., 2008. [Online]. Available: http://www.bcs.org/upload/pdf/data-centre-energy.pdf

[29] A. Berl *et al.*, "Energy-efficient cloud computing," *Comput. J.*, vol. 53, no. 7, pp. 1045–1051, 2010.

[30] E. Gelenbe, "Energy packet networks: Adaptive energy management for the cloud," in *Proc. CloudCP '12 Proc. 2nd Int. Workshop Cloud Comput. Platforms*. ACM, 2012, pp. 1–15. [Online]. Available: https://doi.org/10.1587/nolta.9.322

[31] E. Gelenbe, "Energy packet networks: Smart electricity storage to meet surges in demand," in *Proc. 5th Int. Conf. Simul. Tools Techn.*, 2012, pp. 1–7.

[32] J. M. Fourneau, A. Marin, and S. Balsamo, "Modeling energy packets networks in the presence of failures," in *Proc. 24th IEEE Int. Symp. Model., Anal. Simul. Comput. Telecommun. Syst.,* 2016, pp. 144–153.

[33] E. Gelenbe, "G-networks by triggered customer movement," *J. Appl. Probab.*, vol. 30, no. 3, pp. 742–748, 1993.

[34] E. Gelenbe and A. Stafylopatis, "Global behavior of homogeneous random neural systems," *Appl. Math. Model.*, vol. 15, no. 10, pp. 534–541, 1991.

[35] E. Gelenbe, "Steady-state solution of probabilistic gene regulatory networks," *Phys. Rev. E*, vol. 76, no. 3, 2007, Art. no. 031903.

[36] E. Gelenbe and O. H. Abdelrahman, "An energy packet network model for mobile networks with energy harvesting," *Nonlinear Theory Appl., IEICE*, vol. 9, no. 3, pp. 322–336, 2018.

[37] E. Gelenbe and C. Morfopoulou, "A framework for energy-aware routing in packet networks," *Comput. J.*, vol. 54, no. 6, pp. 850–859, 2010.

[38] E. Gelenbe and C. Morfopoulou, "Power savings in packet networks via optimised routing," *Mobile Netw. Appl.*, vol. 17, no. 1, pp. 152–159, 2012.

[39] E. Gelenbe and T. Mahmoodi, "Energy-aware routing in the cognitive packet network," in *Proc. ENERGY 2011, 1st Int. Conf. Smart Grids, Green Commun. IT Energy-Aware Technol.*, 2011, pp. 7–12.

[40] E. Gelenbe and E. T. Ceran, "Central or distributed energy storage for processors with energy harvesting," in *Proc. Sustain. Internet ICT for Sustainability (SustainIT)*, 2015, pp. 1–3.

[41] E. Gelenbe and E. T. Ceran, "Energy packet networks with energy harvesting," *IEEE Access*, vol. 4, pp. 1321–1331, 2016.

[42] Y. Yin, "Optimum energy for energy packet networks," *Probab. Eng. Informational Sci.*, vol. 31, no. 4, pp. 516–539, 2017.

[43] Y. M. Kadioglu, "Finite capacity energy packet networks," *Probab. Eng. Informational Sci.*, vol. 31, no. 4, pp. 477–504, 2017.

[44] Y. M. Kadioglu and E. Gelenbe, "Product-form solution for cascade networks with intermittent energy," *IEEE Syst. J.*, vol. 13, no. 1, pp. 918–927, Mar. 2018. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8424547

[45] E. Gelenbe, "Synchronising energy harvesting and data packets in a wireless sensor," *Energies*, vol. 8, no. 1, pp. 356–369, 2015. [Online]. Available: https://doi.org/10.3390/en8010356

[46] E. Gelenbe and A. Marin, "Interconnected wireless sensors with energy harvesting," in *Proc. Anal. Stoch. Model. Techn. Appl.—22nd Int. Conf. Proc.*, 2015, pp. 87–99.

[47] R. Takahashi, T. Takuno, and T. Hikihara, "Estimation of power packet transfer properties on indoor power line channel," *Energies*, vol. 5, no. 7, pp. 2141–2149, 2012.

[48] R. Takahashi, S.-I. Azuma, K. Tashiro, and T. Hikihara, "Design and experimental verification of power packet generation system for power packet dispatching system," in *Proc. Amer. Control Conf.*, 2013, pp. 4368–4373.

[49] E. Gelenbe and K. C. Sevcik, "Analysis of update synchronization for multiple copy data bases," *IEEE Trans. Comput.*, vol. C-28, no. 10, pp. 737–747, Oct. 1979.

[50] E. Gelenbe, "Réseaux neuronaux aléatoires stables," *Comptes Rendus de l'Académie des Sci. Srie 2*, vol. 310, no. 3, pp. 177–180, 1990.

[51] P. G. Harrison, "G-networks with propagating resets via RCAT," *SIGMETRICS Perform. Eval. Rev.*, vol. 31, no. 2, pp. 3–5, 2003. [Online]. Available: http://doi.acm.org/10.1145/959143.959144

[52] E. Gelenbe, "Dealing with software viruses: A biological paradigm," *Inf. Secur.* Tech. Rep., vol. 12, no. 4, pp. 242–250, 2007.

[53] J. M. Fourneau, K. Wolter, P. Reinecke, T. Krauß, and A. Danilkina, "Multiple class g-networks with restart," in *Proc. ACM/SPEC Int. Conf. Perform. Eng.*, Prague, Czech Republic, Apr. 21–24, 2013, pp. 39–50. [Online]. Available: http://doi.acm.org/10.1145/2479871.2479880

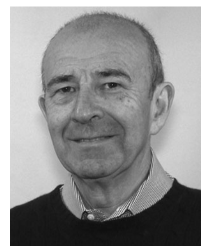
[54] J. M. Fourneau and K. Wolter, "Some applications of multiple classes g-networks with restart," in *Proc. Comput. Inf. Sci.—31st Int. Symp.* Kraków, Poland, Oct. 27–28, 2016, pp. 126–133. [Online]. Available: https://doi.org/10.1007/978-3-319-47217-1_14

[55] J. M. Fourneau, "Mean value analysis of closed g-networks with signals," in *Proc. Comput. Perform. Eng.—15th Eur. Workshop*, Paris, France, Oct. 29–30, 2018, pp. 46–61. [Online]. Available: https://doi.org/10.1007/978-3-030-02227-3_4

[56] E. Gelenbe, "G-networks with signals and batch removal," *Probab. Eng. Informational Sci.*, vol. 7, no. 3, pp. 335–342, 1993.

[57] E. Gelenbe and A. Labed, "G-networks with multiple classes of signals and positive customers," *Eur. J. Oper. Res.*, vol. 108, pp. 293–305, 1998.

[58] E. Gelenbe and J.-M. Fourneau, "G-networks with resets," *Perform. Eval.*, vol. 49, no. 1, pp. 179–191, 2002.



Erol Gelenbe (F'86) received the B.S. degree from Middle East Technical University, Ankara, Turkey, the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of New York University, Brooklyn, NY, USA, and the Doctorat d'État ès Sciences Mathématiques degree from the Université Pierre et Marie Curie, Paris, France.

He is a Professor with the Institute of Theoretical and Applied Informatics, Polish Academy of Sciences (IITIS-PAN), Warsaw, Poland, and a Visiting Professor with the Imperial College, London, U.K. He retired from the Dennis Gabor Professorship at the Imperial College at the beginning of 2019, and from the Professorship at Université Paris-Descartes, in 2005. Renowned for developing mathematical models of computer systems and networks, inventing G-Networks and random neural networks, he has graduated 83 Ph.D. students, including 25 in the U.K. He has been a recipient of honorary doctorates from the Università di Roma II, Italy (1996), Bogaziçi University, Istanbul, Turkey (2004), and Université de Liège, Belgium (2006); he was the recipient of the Parlar Foundation Science Award (1994), the Grand Prix France Télécom of the French Academy of Sciences (1996), the ACM-SIGMETRICS Life-Time Achievement Award (2008), the IET Oliver Lodge Medal (2010), the "In Memoriam Dennis Gabor Prize" of the Hungarian Academy of Sciences (2013), and the Mustafa Prize (2017). Further, he was the recipient of the Knight of the Legion of Honour (2014) and Officer of Merit (2002) of France; and the Commander of Merit (2005) and Grand Officer of the Order of the Star (2007) of Italy, for services to higher education and research. He is on the editorial board of the IEEE TRANSACTIONS ON CLOUD COMPUTING and other journals, and is the Editor-in-Chief of the Springer journals—the *Nature* and the *Computer Science*. He is a Fellow of Academia Europaea (2005), the French National Academy of Technologies (2008), the Science Academy of Turkey (2012); he is a Foreign Fellow of the Royal Academy of Belgium (2015) and the Hungarian (2010) and Polish (2013) Academies of Science. He is also a Fellow of the ACM, the Royal Statistical Society, and the Institution for Engineering and Technology.

Yunxiao Zhang (S'17) received the B.Sc. degree in electrical and electronic engineering from Newcastle University, Newcastle upon Tyne, U.K., and the M.Sc. degree in control systems from the Imperial College London, London, U.K., where he is currently working toward the Ph.D. degree in intelligent systems and networks group, at the Department of Electrical and Electronic Engineering, with Professor Erol Gelenbe.

He has authored or coauthored several conference and journal papers. His research interests lie in energy harvesting wireless networks, energy packet networks, and the modeling and performance analysis of computer systems powered by renewable and intermittent sources of energy.