Panayiotis Kolios, Vasilis Friderikos and Katerina Papadaki

Energy efficient relaying via store-carry and forward within the cell

Article (Accepted version)
(Refereed)

Original citation:
Kolios, Panayiotis, Friderikos, Vasilis and Papadaki, Katerina (2014) Energy efficient relaying via store-carry and forward within the cell. IEEE Transactions on Mobile Computing, 13 (1). pp. 202-215. ISSN 1536-1276

DOI: 10.1109/TMC.2012.233

Energy Efficient Relaying via Store-Carry and Forward within the Cell

Panayiotis Kolios, *Student Member, IEEE,* Vasilis Friderikos, *Member, IEEE,* and Katerina Papadaki, *Member, IEEE*

Abstract—In this paper store-carry and forward (SCF) decision policies for relaying within the cell are developed. The key motivation of SCF relaying stems from the fact that energy consumption levels can be dramatically reduced by capitalizing on the inherent mobility of nodes and the elasticity of Internet applications. More specifically, we show how the actual mobility of relay nodes can be incorporated as an additional resource in the system to achieve savings in the required communication energy levels. To this end, we provide a mathematical programming formulation on the aforementioned problem and find optimal routing and scheduling policies to achieve maximum energy savings. By investigating structural properties of the proposed mathematical program we show that optimal solutions can be computed efficiently in time. The trade-offs between energy and delay in the system are meticulously studied and Pareto efficient curves are derived. Numerical investigations show that the achievable energy gains by judiciously storing and carrying information from mobile relays can grow well above 70% for the macrocell scenario when compared to a baseline multihop wireless relaying scheme that uses shortest path routes to the base station.

Index Terms—Energy consumption, Store-carry and forward relaying, Delay tolerance, Wireless routing, Green cellular networking, Network flows.

1 Introduction

Energy utilization has become imperatively important for network operation since the accumulated effect of increased data rates over cellular networks and volume of data traffic that needs to be supported may lead to unsustainable requirements on energy consumption in wireless networks [1][2][3][4][5][6].

Fixed and/or mobile wireless non-regenerative or regenerative multihop relaying is currently considered as one of the most fundamental architectural elements to be integrated into the next generation of cellular networks [7]. The rational of incorporating multihop relaying within the cell are multi-faceted. The most well documented ones relate to the ability of multihop relaying to enhance cell coverage, the spatial reuse of the scarce wireless resources and user throughput, especially at the cell edge, compared to traditional cellular architectures where mobile nodes transmit directly to the Base Station (BS). A further seminal aspect of multihop architectures, that has more recently attracted research attention, is their potential to reduce the overall energy consumption in the network [8].

Multihop paths benefit from the superlinear relationship of the communication distance to the signal propagation losses. As such, there may exist a path from source to destination where the accumulated (re)transmission energy and the circuit energy consumption of all the participating transceivers en route, consist only a fraction of the required direct transmission energy consumption.

In this work, we present an architectural extension to current cellular network deployments that can potentially achieve many-fold reductions in the end-to-end communication energy consumption. The work is motivated by the need to realize substantial energy efficiency gains in the operation of cellular networks. This latter problem has been exacerbated recently by the growing adoption of mobile Internet over the subscriber base. Evidently, the introduction of data services on mobile phones has caused a considerable rise in the energy consumption of infrastructured systems. On the other hand, and unlike voice services with stringent quality of service requirements, Internet type services vary greatly in content, with very broad delivery delay constraints. Capitalizing on the elasticity of such delay tolerant services, mobile terminals can possibly postpone the transmission of information messages while in transit and only engage in communication at locations within the cell with favorable channel gains. Doing so, both the terminals and the BS require less power to communicate the information and thus conserve energy. Hence, in contrast to previous studies, we propose in this work *mechanical relaying* where mobile nodes store and carry information messages while in transit and only forward the data at a later time when they experience better channel conditions. Such a store-carry and forward relying scheme can be utilized by a plethora of different elastic Internet applications such as for example email access, FTP and Peer-to-Peer file transfers, RSS feeds, status updates from social networking sites and over-

- *P. Kolios and V. Friderikos are with the Center for Telecommunications Research, King's College London, London, U.K., WC2R 2LS. E-mail: {panayiotis.kolios, vasilis.friderikos}@kcl.ac.uk*
- *K. Papadaki is with the Operational Research Group, London School of Economics, London, U.K., WC2A 2AE. E-mail: k.p.papadaki@lse.ac.uk*

the-air software updates for the terminal.

It is important to note at this point that the proposed scheme differs in at least two major and important ways to the store-carry and forwarding scheme already proposed in the context of delay tolerant networks (DTNs). Firstly, while in DTNs storing and carrying the message is an inevitable need in order to provide connectivity between communicating parties, here connectivity is not an issue. Infrastructure nodes, i.e., BSs, provide an almost 100% connectivity in cellular networks and instantaneous communication can be established. However transmissions are deliberately delayed to target possible energy efficiency gains. Secondly, in DTNs due to the unexpected and frequent disconnections, protocols replicate messages at node encounters to increase the delivery success rate. On the contrary here no message replication is necessitated as mobile terminals can communicate directly with the BS and thus message delivery is guaranteed.

Finally the proposed scheme differs from the already proposed basic multihop schemes where messages are forwarded as soon as they become available at the relay node. Hence, the proposed message forwarding scheme can be considered as an *all-inclusive* multihop relaying mechanism where both the traditional direct single hop and basic multihop routing schemes can be employed. In addition however, store-carry and forwarding can be realized.

We illustrate the proposed relaying extension in figure 1. We identify three communicating network entities: (i) source nodes (or user terminals (UT)), (ii) relay nodes (in this case relays are assumed to be communicating radios on board vehicles and thus the abbreviation Vehicular Relay (VR) is used) and (iii) centralized base stations (BSs). In the traditional cellular networks, all source nodes communicate with the base station (BS) directly. In the basic multihop approach, message forwarding takes places as soon as the information is received by the relay nodes. In SCF relaying however, nodes have the flexibility to postpone message forwarding while in transit and thus have the potential to achieve localized transmissions, for example when passing by the BS. Figure 1 illustrates the possible message forwarding links for the uplink case. Note that in the same manner the downlink can benefit from the message postponement scheme where a BS defers transmissions for preferable future time instances. Moreover, there are a plethora of incentives for mobile nodes, including economic drives, that could motivate adoption of the scheme. Even though this is not part of this paper, incentivizing store, carry and forward relaying of information messages is an issue that needs to be addressed. The former issue is addressed in [9], where the authors investigate auctioning strategies for relaying in green cellular networks. Importantly, such techniques can be extended for the SCF relaying paradigm as well.

The paper contributions are summarized as follows:

1) A store-carry and forwarding scheme is introduced for cellular networks.
2) A theoretical and practical networking model is derived and in the sequel it is demonstrated that this model can be applied for multi-user, multi-rate problems with dynamic node mobility patterns for both the uplink and downlink scenarios. The model includes and utilizes the traditional single hop and the basic multihop scheme and further benefits from store-carry and forward relaying.
3) An innovative network flow formulation is derived where the optimal routing and scheduling policies are constructed in small running times even for large network instances.

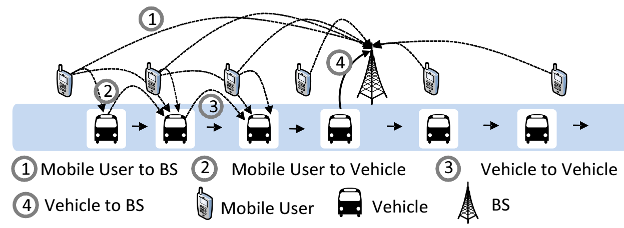


Fig. 1: Different multihop relaying realizations within the cell (including store-carrying and forwarding).

An overview of the mechanical relaying concept can be found in [10] while the work in [11] details the intercell interference reduction gains of the proposed architecture. The case of multi-cell scenarios, where elastic traffic can be delivered to nearby BSs in order to achieve a trade-off of energy, delay and load-balancing between the cooperative cells, is developed and studied in [12]. Fixed and/or mobile relaying has been considered in several cases in the past to address load imbalance problems caused by the very nature of cellular architectures. Relay nodes provide alternative paths for data exchange via the support of the neighbouring BSs (see for example the work in [13]). The benefit of MR over traditional approaches however is that given adequate delay tolerance, a mobile node can store and carry information while in transit to an adjacent cell before forwarding takes place. In this way, high energy efficiency gains can be maintained while achieving the desired load-balancing effect.

Nevertheless, in this paper we focus on the single cell case, detailing the fundamental operations of MR. In contrast to all previous work, here we examine in detail the proposed green networking architecture for energy efficiency, elaborate on the different properties of the mathematical framework and delve into a thorough investigation on the optimization characteristics of the proposed relaying paradigm. More specifically, detail studies are presented to illustrate the flexibility of the proposed mathematical model for a plethora of underlying mobility conditions and the practicality of the derived mathematical programming problem is illustrated for the general multi-user, multi-rate problems.

The rest of the paper is organized as follows. In Section 2 we describe the network model under investigation and in Section 3 we formulate a mathematical program for the proposed strategy. Moreover, in Section 4 we provide numerical investigations on the performance of the proposed scheme. Section 5 dwells upon the effects of the model structure on the system performance. In Section 6 we study the multirate extension of the problem formulated in section 3. Finally, in Section 7 we position our work to the current research activity and conclude the paper in Section 8.

2 System Architecture

To begin with, we consider the uplink of information from all active terminals to the BS but the analysis is valid for the downlink case as well. We consider a BS within an one-dimensional cell of radius R. The one-dimensional realization is made to aid visualization in the forthcoming figures and assumed to model a road segment. However, we relax this assumption in subsequent sections of the paper and study a more realistic cell topology (see section 5.3). The BS collects location and mobility information for all nodes in the system and is responsible for deriving the end-to-end forwarding paths. The work in [14][15][16] considers how such information can be collected and shared among participating nodes efficiently. Further, the BS serves a number of nodes that are either static or mobile. To distinguish among source nodes and relay nodes, we assume here that all source nodes are static while the rest are mobile. This assumption is done to show that static (pedestrian) users can be considered by the model as well. However the model described below is applicable for the more general case where all nodes can be assumed as mobile.

Based on the above, we consider hereafter the set $\mathcal{M} = \{1, ..., M\}$ of active users (UTs) that are static and the set $\mathcal{N} = \{1, ..., N\}$ of mobile nodes that are either dedicated relay entities or idle user terminals inside moving nodes. The mobile nodes considered here can be vehicles in a road network. It is further assumed that all vehicular relays (VRs) travel within a road network with velocity $v_i,\ i \in \mathcal{N}$.

2.1 Communications Model

For the communication model we assume that all nodes in the system can transmit/(receive) to/(from) a single other node at any one time. All UTs have a single message of size F (bits) to send to the BS and a data rate of B (bps) is supported by all links. We relax the latter two assumptions in the sequel (see sections 5.4 and 6.1 respectively). Moreover, we identify three sources of energy consumption during communication. The energy consumed by the transmitter electronics, e_t, the circuit energy consumption at the receiver side, e_r and the energy consumed by the power amplifier e_d at the transmitter side.

Furthermore, distance depended signal attenuation is considered. For line-of-sight paths (within the range d_{brake}) free space losses are assumed whereas for non-line-of-sight paths the two-ray model is used.

In this dynamic and time varying network, two communicating entities i, j have an instantaneous relative velocity $\bar{v}_{ij}$. Let D_{ij} be the initial distance between the transceiver pair i, j. Then $g(D_{ij}, t)$ is the absolute distance at time t between nodes i, j during transmission, which is given by:

$$g(D_{ij}, t) = |D_{ij} - \bar{v}_{ij} t|. \qquad (1)$$

The total energy consumption in transmission of a single message is given by:

$$f_{ij}(s) = (e_r + e_t)Bs + Be_d \int_0^s g(D_{ij}, t)^\eta dt, \qquad (2)$$

where s is the communication time and η is the pathloss exponent.

2.2 Space-Time Network

To generate feasible paths for message routing, knowledge is needed on the future positions of mobile nodes. To capture the dynamics of the network, node positions are replicated at consecutive time epochs of length $\tau = \frac{F}{B}$ units of time over the time horizon $\mathcal{T}$. In this way a *space-time network* is generated over the interval of time $\mathcal{T}$ as shown in figure 2. In the illustration one UT, four VRs at three consecutive time intervals and a single BS are shown. To aid visualization, on this figure all vehicles travel in the same direction. This does not need to be the case in general for the model to be valid. Individual vehicle movement can vary in speed and direction.

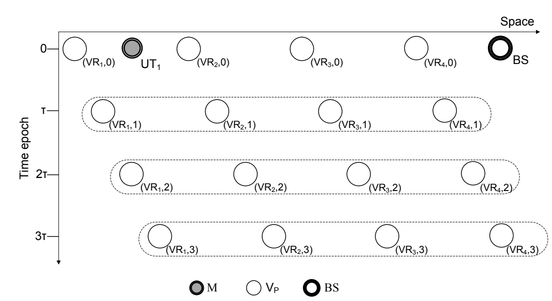


Fig. 2: Time expanded network with mobile node replication at consecutive time epochs. In this way a static network is obtained over the time horizon $\mathcal{T}$.

The nodes in this space-time network are defined as follows: The entry (k, a) uniquely identifies the position of vehicle k at time epoch a and the set V_p contains all such vehicle positions in the interval $\mathcal{T}$. Therefore, node $(VR_1, 0)$ determines the current position of relay 1 while node $(VR_1, 1)$ defines the position of VR_1 at time epoch 1, and so on. Since all UTs and the BS are static entities, their distance does not change with time and therefore can be uniquely identified by their indexed value.

The basic principles of space-time networks have been extensively used over the last fifty years with significant practical use in transportation and logistics. More relevant to our study is their use in DTNs. Of notable importance is the work in [17] where the authors consider the space-time model for sparse networks and suggest the store-carry and forward technique for routing messages. In a similar manner, the work in [18] considers the space-time network with deterministic mobility such that the contact times between nodes can be estimated.

The clear distinction we try to make here from all previous works in the area of DTNs is that for our case, connectivity is not an issue as all nodes can communicate directly with the BS and informed routing decisions can be made. Therefore the space-time network under consideration is not restricted by the contact opportunities as is the case with DTNs. Further, while the above referenced works have studied various different objectives (minimizing delay or maximizing bandwidth under energy constraints), the optimization model under investigation is significantly different (investigating the tradeoff between energy and delay).

3 Mathematical Model

In this section we construct the mathematical model of the proposed SCF network architecture. To ease interpretation, the model is first described for a single UT and the multi-user scenario is explained later (see section 3.2). We define the space-time network for a single UT as a directed graph $G = (V, L)$ such that $V = V_p \cup (UT, BS)$ and L is the set of links. The links in the network represent possible movement of information messages in space and time. We partition links L into two sets:

- L_1 : set of links that represent transmission of information from one transceiver to another. Communication can be between vehicles, the UT and the BS.
- L_2 : set of links that carry flow via vehicle movement. In this case no transmission occurs.

Links L_1 are the set of links defined as follows:

$$\begin{array}{llll} L_1^A: & (k,a) & \mapsto & (l,a+1) \\ L_1^B: & UT & \mapsto & BS \\ L_1^C: & UT & \mapsto & (k,a+1) \\ L_1^D: & (k,a) & \mapsto & BS \\ \multicolumn{4}{l}{\text{for } k,l \in \mathcal{N}} \\ \multicolumn{4}{l}{\quad (k,a),(k,a+1),(l,a+1) \in V_p.} \end{array} \tag{3}$$

Links L_1^A are the set of links that establish communication between vehicular relays k and l. Links in the set L_1^B are the direct single hop communication between a UT and the BS. Note that both entities in L_1^B links are static and thus no message postponement occurs. Links L_1^C are the set of links between the UT and relay k. In this case, the tail node UT, does not have a time component and thus the communication delay is defined by the head node of the link. Illustratively, for link $UT_1 \mapsto (VR_2, 2)$ in figure 3, the user terminal postpones transmission for a single time interval and initiates transmission with vehicle 2 at time $t = \tau$ and ends transmission at time $t = 2\tau$. Finally, links L_1^D are the set of communication links between vehicle k and the BS.

Links L_2, defined as $L_2 = (k,a) \mapsto (k, a+1) : \forall (k,a),(k,a+1) \in V_p$, are those links carrying flow by vehicle movement at consecutive positions of vehicle k. An example of a link in L_2 can be seen in figure 3 for $(VR_2, 2) \mapsto (VR_2, 3)$.

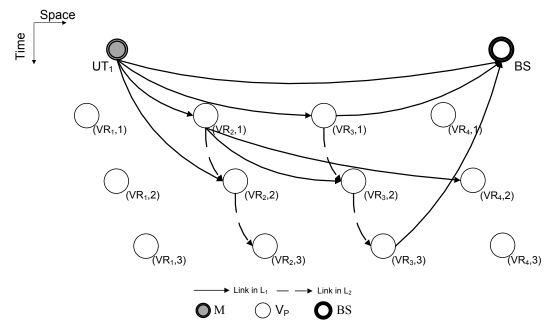


Fig. 3: Space-time network for proposed network architecture. Selectively a number of links are shown.

In the space-time network, the directed paths between nodes UT and BS represent all possible (sensible) ways of getting a message from the UT to the BS either in a single hop or multihop fashion.

We define $x_{ij} \in \{0, 1\}$ to be the flow on link $i \mapsto j \in L$ with capacity $u_{ij} = 1$. Further, since we want one unit of flow to leave the UT and reach the BS, we define the supply and demand for all nodes $i \in V$ as follows:

$$b(i) = \begin{cases} +1 & \text{i = UT} \\ -1 & \text{i = BS} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

3.1 Energy and Delay cost

The energy consumed in transmission of a unit flow between a transceiver pair $i,\ j\ \in V$, is as follows:

$$E(i \mapsto j) = \begin{cases} f_{ij}(\tau) & for\ \ i \mapsto j \in L_1 \\ 0 & for\ \ i \mapsto j \in L_2, \end{cases} \tag{5}$$

where $f_{ij}(\tau)$ is defined in section 2. It must be noted that for links $i \mapsto j \in L_1^A \cup L_1^B$ with $\bar{v} = 0$, energy consumption is proportional to the flow x_{ij} with a multiplier of $f_{ij}(\tau)$. This is not the case for links in $L_1^C \cup L_1^D$ with $\bar{v} \neq 0$. However, since x_{ij} takes binary values, $x_{ij} f_{ij}(\tau)$ gives the correct energy consumption, since $f_{ij}(\tau)$ is the amount of energy used for one unit of flow. The delay per unit flow for all links in L is:

$$\Phi(i \mapsto j) = \begin{cases} \tau & for\ \ i \mapsto j \in L_1^A, L_1^B, L_1^D, L_2 \\ \tau(a+1) & for\ \ i \mapsto j \in L_1^C, \end{cases} \tag{6}$$

where for L_1^C links and $j = (k, a+1)$, the user postpones transmission for a time periods before transmitting. We define the total cost per unit flow as a weighted sum of the energy consumed in transmission and the delay incurred while traversing the link,

$$c_{ij} = E(i \mapsto j) + \lambda \Phi(i \mapsto j). \tag{7}$$

where λ is a weighting parameter of the delay cost for information to traverse the link $i \mapsto j$. Note that for the case where hard deadline constraints need to be imposed, the time horizon $\mathcal{T}$ for node replication can be limited to the maximum delay tolerance of the services considered. All other model derivations discussed above remain the same. For the latter case, the cost for all links in L_2 would be zero and all links in L_1 will have an energy cost term only.

3.2 Mathematical model for Multiple Users

The above space-time network can be generalized for M users in the set $\mathcal{M}$, by replacing the single UT with the elements of $\mathcal{M}$ in the set of nodes and links. Links in L_1 were defined for binary values of the flow and unit capacity. We would like to keep the flow binary for links in L_1, since it is also needed for the energy consumption cost to be correct, and keep the flow integer for links in L_2. Integrality of the flow is a natural assumption, since we are transporting messages, and we show that restricting the flow to at most one unit for links in L_1 does not cause any problems in the network with M users:

1) Transmission links of the form $UT \mapsto j$, for some user UT and some node $j \in V_p \cup \{BS\}$, cannot have flow exceeding one unit since each user has a supply of one unit.
2) However, transmission links of the form $(k, a) \mapsto j$, for some vehicle node $(k, a) \in V_p$ and some node $j \in V_p \cup \{BS\}$, might have more than one unit of flow to transmit that initiated from different users.

Case 1) above does not pose a problem and case 2) is resolved in the following proposition (for proof see appendix A):

Proposition 1: For each transmission link $(k, a) \mapsto j$, where (k, a) is a mobile node and $j \in V_p \cup \{BS\}$, that transmits m units of flow, there exist m distinct paths of unit flow from (k, a) to j that have the same cost.

Thus, in the M-user space-time network we can define all links in L_1 to have unit capacity.

We define the capacities for the space-time network of M users as follows: $u_{ij} = 1$ for L_1 links and $u_{ij} = \infty$ for L_2 links. Note that we place no restrictions in the capacity of L_2 links as we simply assume here that large memory availability is not a constraint - but a requirement, in current and future mobile phone manufacturing. However, for the case whereby a study on the effect of memory constraints needs to be undertaken, the capacity of those links (i.e., $u_{ij}, i \mapsto j \in L_2$) could be set accordingly.

Finally, each user $i \in \mathcal{M}$ will have a supply of $b(i) = +1$ and the BS will have a demand of M, $b(BS) = -M$.

3.3 Mathematical programming formulation

In this section we define the joint routing and scheduling problem for the case of M active users and N mobile relays in a single cell that emerges from the discussions above. The optimal forwarding policies can be found by solving problem (P1) below which is a linear integer mathematical program.

The objective function (8) minimizes the cost of all users (source nodes) forwarding their message to the BS (sink node) via any path in the network. Constraints (9) and (10) ensure that radios do not transmit to more than one radio and do not receive information from more than one radio at any time. Constraints (11) are the flow conservation constraints that create paths from the users to the BS to ensure that the supplies/demands are satisfied.

$$\text{(P1)} \min \sum_{i \mapsto j \in L} c_{ij} x_{ij} \tag{8}$$

$$\text{s.t.} \sum_{j: i \mapsto j \in L_1} x_{ij} \leq 1 \;\; \forall\, i, \;\; i \in V_p \tag{9}$$

$$\sum_{k: k \mapsto i \in L_1} x_{ki} \leq 1 \;\; \forall\, i, \;\; i \in V_p \tag{10}$$

$$\sum_{j: i \mapsto j \in L} x_{ij} - \sum_{k: k \mapsto i \in L} x_{ki} = b(i) \;\; \forall\, i, \;\; i \in V \tag{11}$$

$$0 \leq x_{ij} \leq u_{ij}, \;\; x_{ij} \in \mathbb{Z} \tag{12}$$

Solving the above mathematical program using integer programming is hard. However, we show that by relaxing the integrality constraints of the above problem integer solutions are obtained.

Theorem 1: Let $Ax \leq q$ be the matrix representation of constraints (9)-(12), ignoring integrality constraints. Then A is totally unimodular (TU).

The proof is detailed in appendix B. It is well known that if A is TU and q is integer then the feasible region described by $Ax \leq q$ has integer extreme points. Thus, we can relax the integrality constraints and still get integer solutions using linear programming (simplex method). This means that the above mathematical program can be solved for a large number of nodes in very short running times.

4 Numerical Investigations

By solving the mathematical program in section 3.3 to optimality, energy-efficient policies are obtained for different delay requirements in message delivery. We study the micro-cell scenario of radius $R = 400$m with $M = 10$ uniformly distributed active users and $N = 20$ relay nodes. For the communication model we estimate that the distance for switching between the two modes of propagation is $d_{brake} = \frac{4\pi h_t h_r}{\nu}$ where $h_t = 1.5$m and $h_r = 1.5$m are the transmit and receive antenna heights respectively and the transmit radio wavelength is $\nu = 0.126$m. For the two signal propagation modes the following models are used: Free space losses are assumed for communication distances less than d_{brake} with path loss exponent $\eta = 2$ while the plane earth model is used for larger distances with path loss exponent

$\eta = 4$. The fixed circuit power to operate the transceiver electronics is $P_c = 50$mW and the received power threshold for successful communication is $P_r = -52$dBm[1]. Based on these values the following parameters can be deduced: The energy per bit consumed by the power amplifier is $e_d(los) = \frac{P_r(4\pi)^2}{B\nu^2}$ for free space losses and $e_d(mp) = \frac{P_r}{Bh_t^2 h_r^2}$ otherwise. The fixed circuit energy consumption is $e_t = e_r = \frac{P_c}{B}$.

All users have a single message of $F = 4$Mbit to transmit to the BS and the data rate for all links is $B = 1$Mbps while all communications are restricted to the range of R meters. To provide an insight on the performance gains of the proposed scheme we study a network topology with all nodes lying along a single stretch of a road. On this topology, we first consider vehicles traveling in a single direction and bi-directional traffic is included later when a real network topology is used. SUMO[2] mobility simulator is used to generate realistic mobility traces for all relay nodes. SUMO is a microscopic road traffic simulator that captures the mobility of individual nodes within a road network and it is based on the car following principle of vehicular traffic modelling. The car-following parameters used in this work for mobile relay nodes are as follows: maximum acceleration is 0.8m/s^2, maximum deceleration is 4.5m/s^2, maximum travelling speed 14m/s, the car length is 5 meters and response time to unpredictable events is set to 0.5 seconds.

This section first considers the case where vehicles travel at a constant average speed along the roadway and their initial location is uniformly distributed across the cell. Clearly, in real scenarios inter-vehicle interactions will occur depending on the actual distance between vehicles. In section 5.2 we investigate how this inter-vehicle interactions can affect the computed message forwarding decisions by modelling the uncertainty in vehicle positions using SUMO's mobility prediction model. Therefore, in this section we assume that perfect knowledge on future vehicle positions is available and the latter assumption is relaxed in section 5.2.

We compare the proposed mechanical relaying strategies with a baseline multihop (MH) scheme where messages are transmitted as soon as they are received (in a decode and forward manner). Hence, in the latter case we solve problem (P1) with the restriction that no flow can pass through L_2 links. Figure 4 plots the energy-delay trade-offs achieved via the traditional cellular communication, the basic MH and the proposed SCF schemes. As expected, for tight deadlines SCF does not utilize the mobility of nodes to forward information and only utilizes transmission links. Hence, the performance of both MH and SCF are identical, as shown in the figure. On the other hand, for elastic deadlines, SCF achieves considerably lower energy consumption levels

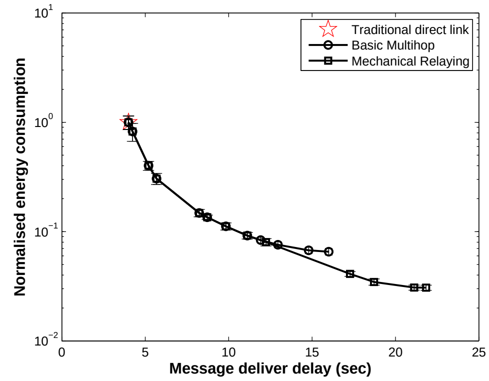


Fig. 4: Energy-delay tradeoff of the direct link, basic multihop scheme and the proposed SCF mechanism.

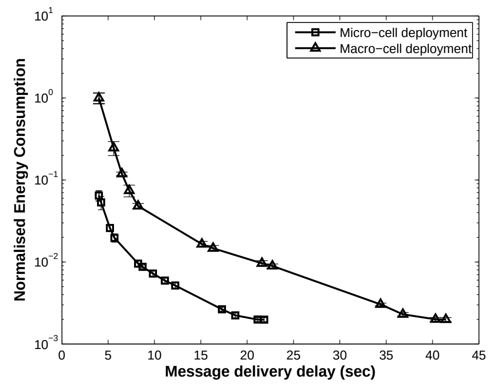


Fig. 5: Pareto energy-delay curves for two cell size deployments.

compared to the basic multihop scheme by benefiting form the longer message delivery delays. Looking into the results in figure 4 it is evident that MH cannot obtain further savings in energy consumption for delays greater than 16 seconds, while SCF takes advantage of higher delay levels to achieve lower energy consumption levels by simply postponing communication for future time instances. Interestingly, for the same limiting delay, the proposed SCF scheme can achieve less energy consumption levels compared to MH. Therefore, SCF provides an enhanced performance even compared to the basic multihop scheme, in addition to the gains observed from the direct transmission link. This advantage of SCF over MH will also be address in section 5.3.

It must be noted that the multihop strategy in comparison differs from that specified in the upcoming 802.16j [20] and LTE-Advanced [21] standards where individual packets are forwarded. Here information is forwarded only when the entire message has been received.

Clearly for increasing cell sizes the gains of the proposed SCF scheme increase compared to the single hop and basic multihop schemes. Figure 5 compares the schemes for two cell size deployments; a micro-cell with $R = 400$m and a macro-cell with $R = 800$m. For elastic services with loose message delivery deadlines, the SCF scheme achieves the same minimum system energy consumption per user in the cell for both cell sizes. From the data values it can be seen that gains

1. This is similar to the average signal strength for an LTE receiver at 10MHz channel bandwidth which is calculated in [19] to be -56.5dBm.
2. http://sumo.sourceforge.net/

well above 50% and 70% are observed compared to the basic multihop scheme for the two cell size deployments respectively while compared to the direct-link, many-fold savings in energy can be achieved.

5 Effect of Model Structure on System Performance

In this section a detailed investigation on the effect of the modeling assumptions made in the derivation of the problem formulation is conducted. We show that the parameters of the model structure can be easily adjusted to provide for the maximum energy efficiency gains while avoiding unnecessary computations. For that reason in section 5.1 relay selection algorithms are derived that aim to select a subset of the available mobile nodes as candidate relays for message forwarding. Minimizing the number of candidate relays considered, reduces the overhead and signalling requirements for location updates. Further in section 5.2 the effects of the accuracy in mobility prediction are investigated. Clearly, a single solution to problem (P1) would have given the exact decision policies for the case when mobility of nodes would have been predictable. However, iteratively updating the forwarding decisions at consecutive time periods can combat the uncertainty that arises in scenarios where node mobility is highly unpredictable.

In section 5.3 a real network topology is considered where the strength of SCF to the basic MH scheme is illustrated. Finally the trade-offs in the forwarding decisions accuracy and the increased computation effort are investigated in section 5.4. While considering smaller message sizes reduces τ and allows for increased granularity in the feasible forwarding paths, the number of nodes in the network increases and computational efficiency reduces. As a remark however, for elastic services the effects of this trade-off is significantly less detrimental as shown in the sequel.

5.1 Selection process for a subset of vehicles

The diminishing returns in energy consumption of considering increasingly many vehicles in the system are shown in figures 6 for both soft and hard delay deadlines. In both cases, considering greater than M vehicles in the network only slightly improves the solution while complexity and computational effort increases.

We consider a selection process for a subset of vehicles from set $\mathcal{N}$ and provide bounds for the deviation from the solution that considers all candidate relay nodes. In the first case we consider a selection of $\hat{\mathcal{N}}$ vehicles evenly spaced between the furthest user and the BS. In the second case, we choose the subset $\hat{\mathcal{N}}$ of vehicles based on the delay tolerance of messages while minimizing the energy consumption in the system.

5.1.1 Evenly spaced selection of relay vehicles (ESS)

In this simple strategy, irrespective of the delay tolerance of the service and the user distribution in the cell, we

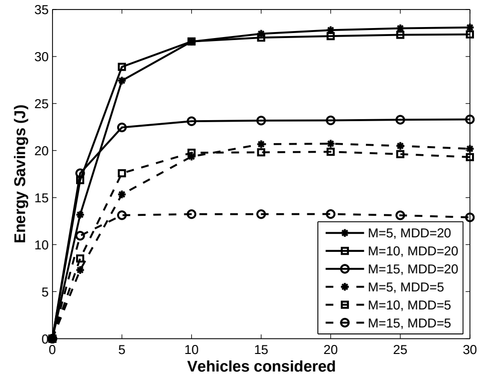


Fig. 6: Diminishing returns in energy savings when considering increasingly many candidate relay nodes. The plots show the returns for both elastic and delay sensitive traffic for message delivery delays (MDDs) of 20 and 5 sec respectively.

select a subset of vehicles evenly spaced between the cell edge and the BS, with a maximum of M vehicles selected. From the set $\mathcal{N}$ we create a subset $\hat{\mathcal{N}}$ of vehicles such that,

$$\hat{\mathcal{N}} = \left\{ i \in \mathcal{N} \colon i = \left\lfloor \frac{N}{M} \right\rfloor \cdot k, \text{ for } k = 1, \ldots, \left\lfloor \frac{N}{\lfloor \frac{N}{M} \rfloor} \right\rfloor \right\} \quad (13)$$

5.1.2 User-vehicle matching (UVM)

Here, the candidate forwarding nodes are selected such that the following expression is minimized:

$$D_{ki} + \lambda \mu_i \quad (14)$$

In (14) the distance D_{ki} is the initial distance of vehicle i from user k and μ_i is an indicator function for selecting a subset of vehicles from the set $\mathcal{N}$. For the minimization program that we are formulating here, this function assigns a small weight to the subset of evenly spaced vehicles and a large weight otherwise. This function can be expressed as follows:

$$\mu_i = \begin{cases} \text{low} & \text{if } \operatorname{rem}(\frac{i}{\lfloor \frac{N}{M} \rfloor}) = 0 \\ \text{high} & \text{otherwise} \end{cases} \quad (15)$$

Therefore, (14) is a weighted sum of the distance (and thus energy consumed) by each user to every mobile node and a factor that tries to evenly space vehicles. Parameter λ as used previously, is the weighting parameter for these two quantities and indicates the importance of message delivery delay to the communication cost. For a selection of $\hat{\mathcal{N}} \subseteq \mathcal{N}$ vehicles the following optimization problem is formulated,

$$\text{minimize} \sum_k \sum_i (D_{ki} + \lambda \mu_i) y_{ki} \quad (16)$$

$$\text{subject to} \sum_k \sum_i y_{ki} = |\hat{\mathcal{N}}| \quad (17)$$

$$y_{ki} \in \{0, 1\} \ \forall\, k \in \mathcal{M},\ i \in \mathcal{N} \quad (18)$$

where y_{ki} is the binary decision variable for selecting vehicle i by user k. We let the cardinality of subset $\hat{\mathcal{N}}$

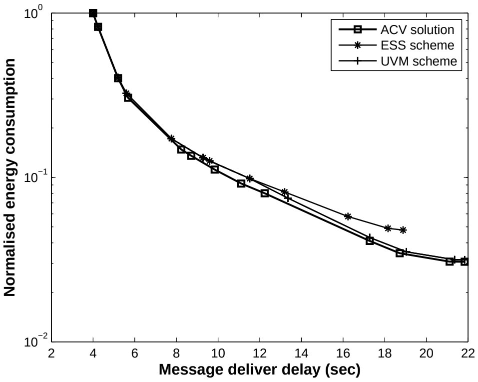


Fig. 7: Energy delay curves for the micro-cell deployment; ACV is the scenario where all candidate vehicles are considered.

equal to the number of users scheduled for transmission, $|\hat{\mathcal{N}}| = M$.

Figure 7 shows the energy-delay curves when selection schemes ESS and UVM are compared to the case where all candidate vehicles are considered (ACV) in a micro-cell deployment scenario. For the results presented here, parameter μ is simply set to 10 and 1000 for the low and high state respectively. From the data values of this figure, ESS deviates from ACV by more than 50% for delay tolerant messages. The UVM method selects vehicles that closely match the user distribution for delay tolerant messages and thus reduces the maximum deviation from ACV to only 16%. Note that for both of these schemes, the size of the subset of vehicles considered is equal to the number of active users in the cell.

5.2 Mobility prediction accuracy

In section 3 we have provided analysis of the problem based on the assumption that future vehicle positions can be predicted accurately by the mobility model. To deal with this problem we introduce an iterative SCF scheme and compared its performance to the following two extreme cases: 1) The case where exact location information for the entire time horizon is available and 2) the case where only initial vehicle location information is available and node replication on the space-time network is done assuming constant speed and direction. The iterative-SCF scheme is described by algorithm 1.

Algorithm 1 Iterative-SCF scheme.

Ensure: k=0.
1: Update vehicle positions based on newly received vehicle location information.
2: Update supply/demand parameters.
3: Re-construct space time network using steps 1, 2.
4: Solve problem (P1) for time horizon $t = k\tau, k\tau + 1, \dots, k\tau + T$.
5: Execute decisions for the first θ time periods.
6: $k = k + \theta$; Go to step 1.

The iterative-SCF scheme as described by algorithm 1 simply re-computes the forwarding decisions at a rate θ. Initially, using the most recent node location information (step 1) and upload requests (step 2), the space-time network is constructed (step 3). Then, the forwarding paths are derived (step 4) for the next T time periods, but only the decisions for the next θ time periods are executed (step 5) before re-computing the forwarding policies.

Figures 8a and 8b illustrate the bounds in the trade-offs achieved by the two extreme solutions for the case of 20 and 30 candidate forwarding relays while figures 8c, 8d show the performance of the iterative-SCF scheme as compared to the optimal solution for the two simulated scenarios. Note that the two simulated scenarios represent the cases where free flow and congested road traffic is experienced by vehicles in a road network.

The effectiveness of re-optimizing the forwarding paths at consecutive time instances is clearly illustrated in figures 8c and 8d. As shown, the deviation factor is kept very small when correcting the forwarding paths in frequent time intervals. Further, with an estimate on the node position over longer time horizons (i.e., T=20 instead of T=10 seconds for both scenarios considered) the deviation factor can be considered negligible. In both figures, a deviation factor of less than 0.2 times from the optimal solution is achieved for T=20 seconds while the forwarding decisions are less deteriorated by the re-computation period θ. Note that re-computing forwarding paths at a rate of θ=4 to θ=8 seconds achieves approximately 30x in energy savings even for highly dynamic topologies (figure 8d).

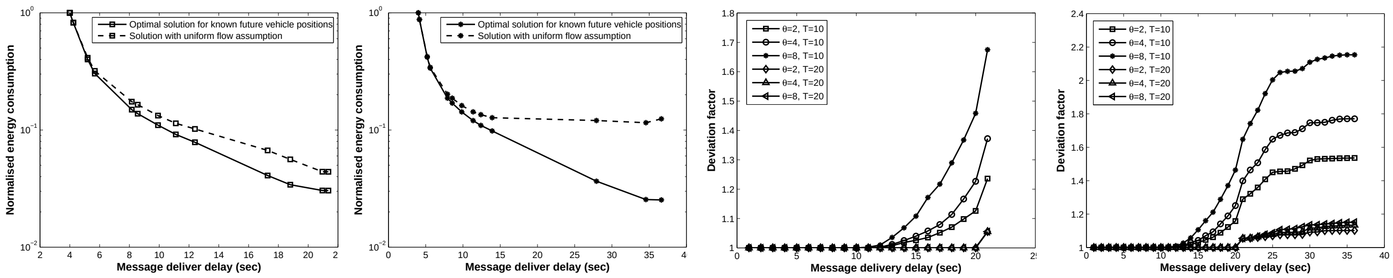
5.3 Experimental evaluation using a real network topology

In the sequel, a realistic network topology is used to evaluate the performance of the proposed SCF scheme and compare it with a baseline MH scheme and the traditional single hop communication. The road network near the Strand Campus of King's College London is investigated (as shown in figure 9). The road network layout is imported from OpenStreetMap[3] into the SUMO simulator and the exact position of a BS node from a UK mobile network provider[4] is used. Note that along the roadway under consideration four traffic light junctions are operating. We evaluate an instance of the problem with a mere $N = 40$ vehicles (in each direction of traffic flow) and $M = 20$ active users along the marked roadway.

The energy-delay performance for the two schemes is shown in figure 10.

As can be seen from this figure, the SCF scheme can achieve the maximum energy savings as opposed to the basic multihop scheme, however at the expense of increase message delivery delay. Nevertheless, its immediately clear from these results that simple multihop

3. http://www.openstreetmap.org/
4. Accurate cellular network deployments and BSs positions can be found in http://www.sitefinder.ofcom.org.uk/.



(a) (b) (c) (d)

Fig. 8: The performance of the two extreme cases with exact mobility prediction accuracy and the uniform flow approximation are illustrated in figures 8a and 8b for 20 and 30 mobile nodes respectively. Further, the performance of the iterative-SCF scheme when compared to the optimal solution is illustrated in figures 8c and 8d for the two simulated scenarios.

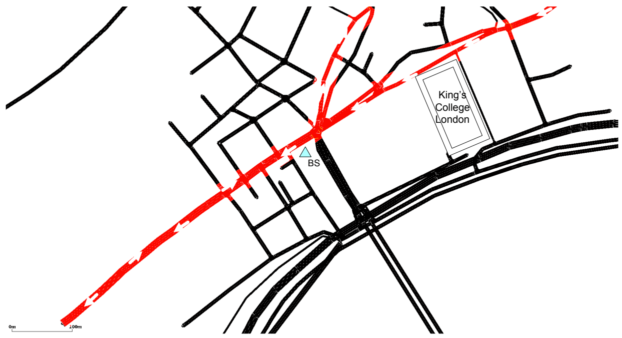


Fig. 9: Network deployment in the area near the Strand Campus at King's College London. A single BS from a UK operator is shown in the figure.

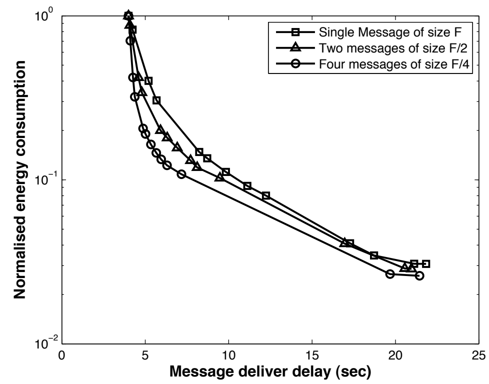


Fig. 11: Energy-delay curves for different packetization levels.

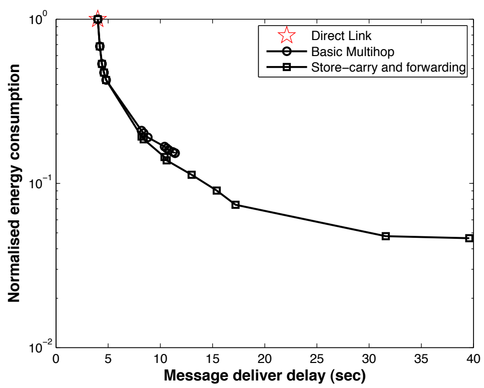


Fig. 10: Energy-delay tradeoffs achieved by the direct link, basic multi-hop and SCF schemes for the network topology in the area near King's College London.

relaying could not suffice as the favorable approach as it fails to take into account future forwarding opportunities.

5.4 Effect of different information message sizes

Traditionally, routing decisions in fixed and mobile networks are performed on a per packet basis to utilize better system resources and respond to network changes. However, the transmission times of packets can be considerably smaller than the vehicular propagation times and thus changes in large scale channel conditions will be negligible. In addition, node replication at such small time intervals could not be realized as the space-time network data set space will increase non-linearly. Note that the number of mobile node positions, $N \cdot \frac{T}{\tau}$ depends on the transmission time and thus on the message size. Nevertheless, reducing the time interval τ increases the number of possible forwarding paths. Figure 11 depicts the case where the same amount of information is transmitted as a single message or smaller chunks of equal size (i.e, 2 messages of size F/2 and 4 messages of size F/4). Improvements on energy savings when smaller packets are considered can be observed for a range of message delivery delays. However, an increase in the number of nodes in the network increases the computational time of problem (P1). For the single message solution in figure 11, the space-time network contains 70 nodes, while 139 and 1200 nodes are used in the 2- and 4-message solutions respectively, resulting in a substantial increase of the problem space. It is clear that a trade-off exists between complexity (and thus computational time) of the solution and the energy reduction improvements from the available paths.

However, note that the gains in considering a higher granularity network are only marginal and thus do not provide significant improvements. This is especially true for elastic services as shown in figure 11.

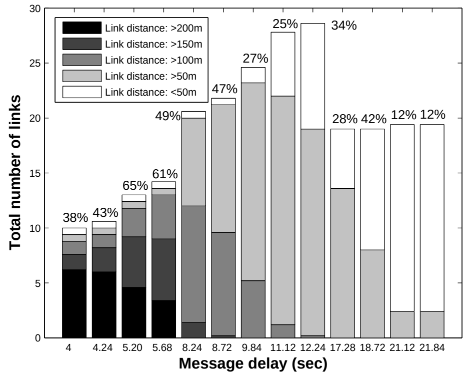


Fig. 12: Communication link distance distribution.

6 Multirate Extension

For the radio communication model described in section 2, the circuit energy consumption e_t and e_r are independent of the communication distance while the transmission energy consumption e_d is a dependent parameter. Moreover, for higher data rates both the receiver sensitivity and signal-to-noise ratio increase and thus the transmission energy consumption e_d increases. However for the fixed circuit energy consumption, e_t and e_r are inversely related to the increased data rate. Therefore for short communication distances where $e_d D^\eta < e_t + e_r$ it may be more energy efficient to transmit at a higher rate, [22][23]. Transmitting at a higher rate reduces the time the transceiver pair is communicating and thus reduces the overall energy consumption.

In figure 12 we provide a breakdown of the communication distance (in steps of 50m) for the optimal forwarding strategies of the micro-cell deployment using the parameters in section 4. As can be seen, there is large variability in the communication distance for different delay constraints. More importantly large variability in the communication distance also exists within each solution (shown by the percentage of links outside the dominating range). These results suggests that further energy savings could be achieved by enlarging the data rate set available for communication.

6.1 Optimal two-rate formulation

In this part we consider the multirate extension of the problem formulated in section 3. We are interested in finding the average consumption gain of the multirate problem with reference to the single rate solution. For simplicity we consider the two rate extension and show how this can be generalized by making the necessary changes in the model and introducing constraints in the formulation of the problem. We thus consider the case where a radio can transmit with rates $\{B_1, B_2\} \in \mathcal{B}$, where $B_1 = 2B_2$. We assume that a unit flow transmitted with rate B_1 takes τ units of time, and 2τ units of time with rate B_2.

As before we consider a space-time network with time steps of duration τ as shown in figure 2. For simplicity we define the two-rate model for a single user but we perform our experiments for multiple users. The links in L_2 remain as defined in the single rate model. There are two types of transmission links in L_1: the ones that transmit with rate B_1, denoted $L_1(B_1)$, that have τ transmission time, and the ones that transmit with rate B_2, denoted $L_1(B_2)$, that have 2τ transmission time. The links in $L_1(B_1)$ are as defined in (3). We define the links in $L_1(B_2)$ as follows:

$$\begin{array}{llll} L_1^E: & (k,a) & \mapsto & (l,a+2) \\ L_1^F: & UT & \mapsto & BS \\ L_1^G: & UT & \mapsto & (k,a+2) \\ L_1^H: & (k,a) & \mapsto & BS \\ \multicolumn{4}{l}{\text{for } k,l \in \mathcal{N}} \\ \multicolumn{4}{l}{\quad (k,a),(k,a+2),(l,a+2) \in V_p.} \end{array} \tag{19}$$

The energy consumption of links in L can then be described as follows,

$$E(i \mapsto j) = \begin{cases} f_{ij}(\tau) & for\ i \mapsto j \in L_1(B_1) \\ f_{ij}(2\tau) & for\ i \mapsto j \in L_1(B_2) \\ 0 & for\ i \mapsto j \in L_2 \end{cases} \tag{20}$$

where the energy consumption function f_{ij} is as defined in equation (2). Similar to the discussion in section 3.1, the energy cost per unit flow takes the correct values when the flow takes binary values. The delay per flow for all links in L is as follows,

$$\Phi(i \mapsto j) = \begin{cases} \tau & for\ i \mapsto j \in L_1^A, L_1^B, L_1^D, L_2 \\ 2\tau & for\ i \mapsto j \in L_1^E, L_1^F, L_1^H \\ \tau(a+1) & for\ i \mapsto j \in L_1^C \\ \tau(a+2) & for\ i \mapsto j \in L_1^G \end{cases} \tag{21}$$

As before, the cost function for all links is a weighted sum of energy and delay characterized by equation (7). We assume $u_{ij} = 1$ for links in L_1, $u_{ij} = \infty$ for links in L_2, $b(UT) = +1$, $b(BS) = -1$ and $b(i) = 0$ for all other nodes. We also introduce some new notation. Let $i = (k,a)$ be a vehicle node of the space-time network. We let $p(i) = (k, a-1)$ denote the *predecessor* of i, and $s(i) = (k, a+1)$ denote the *successor* of i.

The optimization problem is thus expressed with the following mathematical program,

$$\text{(P2)} \qquad \text{minimize} \sum_{i \mapsto j \in L} c_{ij} x_{ij} \tag{22}$$

$$\text{s.t.} \sum_{j: i \mapsto j \in L_1} x_{ij} + \sum_{j: p(i) \mapsto j \in L_1(B_2)} x_{p(i),j} \le 1 \;\; i \in V_p \tag{23}$$

$$\sum_{k: k \mapsto i \in L_1} x_{ki} + \sum_{j: j \mapsto s(i) \in\, L_1(B_2)} x_{j,s(i)} \le 1 \;\; i \in V_p \tag{24}$$

$$\sum_{j: i \mapsto j \in L} x_{ij} - \sum_{k: k \mapsto i \in L} x_{ki} = b(i) \;\forall\, i, \;\; i \in V \tag{25}$$

$$0 \le x_{ij} \le u_{ij}, \;\; x_{ij} \in \mathbb{Z} \tag{26}$$

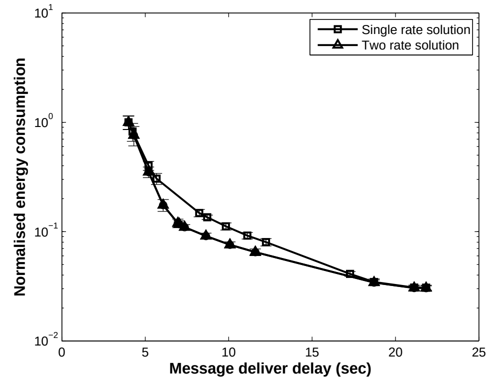
The objective function (22) minimizes the cost of utilizing a path to transfer a message from the user to the BS. The flow conservation constraints of the space-time network are given by (25). Constraints (23) and (24) are the outdegree and indegree constraints respectively. Note that in the outdegree constraints of node $i \in V_p$, we include arcs that emanate from node i and arcs in $L_1(B_2)$ that emanate from node $p(i)$. This is because node $i = (k, a)$ cannot start transmission at time a if there is an active $L_1(B_2)$ arc that emanates from node $p(i) = (k, a-1)$, since vehicle k is still transmitting information at time a. Similarly, in the indegree constraints of node $i \in V_p$, we include arcs that go into i and also arcs in $L_1(B_2)$ that go into $s(i)$. This is because node $i = (k, a)$ cannot start reception at time a if there is an active $L_1(B_2)$ arc that goes into $s(i) = (k, a+1)$, since vehicle k is still receiving information at time a.

Let A_{MR} be the constraint matrix given by constraints (22)-(26). We have the following result:

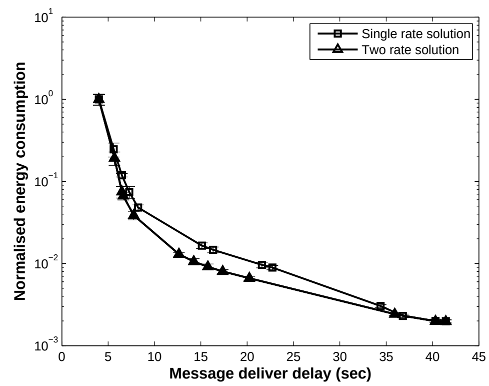
Theorem 2: The matrix A_{MR} is totally unimodular.
The proof is detailed in appendix C. Similar to the single rate model, we can use efficient algorithms to solve the mathematical program described above.

To study the performance of the system we compare the energy consumption gains of the multirate problem solution with reference to the single rate case. Using the same model parameters as in section 4 we solve the mathematical program expressed in equations (22-26) to optimality and compare our results. Note that $B_1 = 1\text{Mbps}$ and for $B_2 = 512\text{Kbps}$ used for this study, the required received power threshold is assumed to be $P_r = -59\text{dBm}$.

We compare in figure 13 the average energy consumption and delay per user in the cell for the solution of the single rate and two rate problem respectively; figure 13a refers to the micro-cell deployment case, while figure 13b relates to the macro-cell deployment. As can be seen in these figures, allowing transceivers to communicated with different rates and optimizing over their selection for a specific rate improves considerably the energy savings. Importantly, significant improvements are achieved for mid-range delays where increased inter-vehicle communication takes place. From the data values, the maximum energy consumption gain of the two-rate solution is 1.72 and 2.05 times for the micro- and macro-scale deployments respectively.



(a)



(b)

Fig. 13: The optimal energy-delay curves for the single rate and two-rate solutions are shown for (a) the micro-cell and (b) macro-cell deployments.

7 Related Work

Energy-efficient communication has recently attracted much attention by both industry and academia. The work in [23] provides a study on the fundamental issues present in energy efficient communications. Further in [24] the authors show that jointly minimizing the fixed circuit energy by hardware sleeping and transmission energy by tuning rate, coding and power, up to 9 times less energy is consumed compared to other adaptive schemes.

Regarding the utilization of node mobility as a resource that can be exploited, the authors in [25] and [26] study the energy efficiency gains that can be achieved in unstructured wireless networks and device simple heuristics to provide good forwarding paths. Under the control mobility of dedicated relay nodes, the work in [27] study energy minimization algorithms to increase the life expectancy of communicating parties.

Research on delay tolerant networks (DTNs) traditionally concentrated on issues of frequent and extended end-to-end disconnections. The work in [28] and [29] provide a detailed study on these issues. In DTNs and for sparsely populated areas, the SCF paradigm has been realized as a viable approach for message forwarding. In similar context to our work, the authors in [30] consider the SCF strategy for message dissemination by exploiting the correlative features of human patterns. Their aim is to minimize the delay in routing messages from source to the destination. In [31], SCF is employed as a fall back mechanism when no end-to-end path exists in vehicular ad-hoc networks. Using mobility prediction, they show that SCF achieves increased packet delivery ratio and reduced data traffic overhead as opposed to other popular routing schemes. To provide increased capacity in a DTN network, the work in [32], [33] consider an architecture where inexpensive, small battery-powered wireless devices (throw-boxes) are deployed to act as relays between the mobile nodes. Algorithms for optimal deployment and energy efficient routing have been developed. However, for all the cases discussed above, the assumption is that end-to-end connectivity is not available. In contrast, in our work forwarding of messages is deliberately delayed for a later time instance when channel conditions are more favorable.

Mobility and opportunistic access to heterogeneous networks is also being currently explored. In [34] an opportunistic threshold-based algorithm (SALSA) is detailed for choosing when to switch between 3G and WiFi infrastructure nodes for information exchange. The referenced paper decides opportunistically on the energy-delay trade-off while our work considers how controlled decision policies can be made by the infrastructure nodes based on knowledge of all the users' position and mobility information. Finally the referenced paper does not capture the possible relaying capabilities of mobile terminals as is the case in the proposed work. [35] is based on incentive principles to tempt user to postpone transmission of delay tolerant traffic for offloading at a later time over a wireless local area network. In this case too (similar to the reference above) the authors assume that several different radio access networks are accessible by the mobile terminals; this is a clear distinction to our study. However, the authors here do not consider energy efficiency but purely the trade-off between user satisfaction and volume of traffic offloading. Also no multihop solutions are provided in the latter work either. Similarly, [36] targets a minimization of cellular traffic via the intentional delaying of data transmissions, targeting future contact opportunities with other mobile nodes that might possess the desired data messages. Such a scenario however is substantially different from the one we consider in this work, since the above referenced paper does not consider the energy efficiency trade-off in delaying data transmissions and further only opportunistic data dissemination is assumed. Here, as exemplified above, we integrate single hop, multihop and delay-tolerant relaying all within an informed decision making process for an energy efficient cellular networking architecture.

In addition to the aforementioned developments, there has been substantial progress on heterogenous technologies for intelligent transportation systems (ITS) by the amalgamation of ad-hoc, cellular and digital broadcast networks, [39][40][41]. For wireless ad-hoc vehicular communications a dedicated access technology is under development, the WAVE suite [42], which approaches with fast pace into standardization. To interact with systems and/or services of a fixed infrastructure and to access the Internet, integration to the cellular network is supported, [43]. In either case, considerable effort has been consumed to address challenges faced in such networks, varying from physical layer channel characterization [44] to position estimation and prediction methods to routing and cross-layer adaptation mechanisms to make the most of the wireless communication interface.

Location information acquisition

Notably, the proposed message forwarding scheme relies on knowledge of the underlying network dynamics to compute the SCF decision policies. As such, location information acquisition could be considered as a primary source of signalling overhead in the realization of mechanical relaying strategies; this is especially true for highly dynamic network topologies where periodic recomputation of the forwarding paths (as described by algorithm 1) is necessary. However, accurate terminal localization has become an increasingly valuable commodity (not only for location based services [45] [46] [47] but also for network optimization [48]) and has been an issue of significant standardization efforts recently, [49] [50]. Not only that, network operators have been increasingly interested in leveraging the terminal capabilities to collect statistics of the underlying network conditions. The document in [51] details the logging and reporting procedures implemented at the user terminals to obtain such information from the surrounding environment. Table 1 below, details the measurement fields standardize by those mechanisms.

TABLE 1: Location information overhead

Parameter	Size (bits)	Definition
Location	63	Lat/ Lon/ Alt information
Time stamp	40	Month/ Day/ Hour/ Min/ Sec
CGI	52	Serving cell id
PCI	288	Neighboring cell id (x32)
Measurements	429	Radio environment measurements

Each log as detailed in the table, contains the position of the user, a timestamp of the event, the serving and neighboring cell serial identification numbers and the received signal strength values. It is important to note here that such information presents a complete set of measurements required for the proposed postponement schemes.

Actual location information can be calculated at the BS for the group of users (via triangulation, trilateration and numerous other techniques as detailed in [45]), computed distributedly at each terminal (via, for example, GPS sensors or triangulation techniques) or in a hybrid fashion by the BS with assistance from the terminals. Of course in the former case, the terminals need not send any signalling to the BS and the additional power cost of location information is negligibly small (as discussed in [45]). The work in [52] illustrates such efficient positioning techniques that achieve a deviation error of less than 25m. For mobility-based solutions, the work in [53] consider how energy efficient and accurate location information can be obtained at the terminal side. Initial location is obtained via GPS receivers while dead reckoning is used to trace node mobility at future time epochs (using the onboard sensors). Through experimentation, the authors show that an accuracy within a 10m error range can be achieved. Finally, the work in [54] consider the case where only cellular information and sensor hints are used to trace terminal motion. As shown by the authors, a mere 5mW is consumed to acquire such information as opposed to the more power hungry GPS alternative. With a constant drain of 20-30mW in standby mode at the terminal, the energy cost of the latter method is clearly negligibly small.

In accordance with the latter work, assume here that location information is calculated at the terminal which then creates a log at consecutive time instances and reports it back to the BS for processing, as described in [51]. Furthermore, let the reporting period equal to the recomputation interval θ as discussed in section 5.2. Then, for $\theta = 8$sec and a delivery delay window of $T = 25$sec, the accumulated log file size (i.e., 3 logs containing the entries in table 1) accounts for 2.5Kbits. Furthermore, assuming that both the source node and an additional neighboring relay (assisting in message forwarding, as shown in section 5.1) send the latter information to the BS, resulting in a total overhead of 5Kbits (clearly, the overhead increases linearly with the number of additional relays observed). Evidently, such control overhead is at least several orders of magnitude smaller than typical file sizes. As such, the overhead requirement in realizing such postponement schemes are only marginally reducing the overall benefits of the proposed store-carry and forwarding paradigm.

8 Conclusions

In this work we address the problem of energy efficiency in cellular networks and propose a store-carry and forward relaying strategy to achieve savings in transmission energy. We show how mobility of relay nodes can be used in a ubiquitous fashion to drastically reduce the transmission energy across different entities in the system. We study how this reduction in energy translates into message delivery delay and obtained optimal tradeoffs for these competitive objectives. We further show attainable performance gains of the proposed strategy that can achieve savings in energy well above 70% compared with a baseline multihop scheme in a macro cellular environment. Importantly, for the same end-to-end delay, the proposed SCF strategy achieves significant energy consumption gains compared to the aforementioned multihop routing. Two simple approximation algorithms are introduced for selecting forwarding nodes in the cell, that utilize the observation that marginal improvements are achieved from considering increasingly many mobile nodes. We have also considered how mobility prediction could prove beneficial in attaining near optimal solutions while maintaining the minimum required communication overhead for location tracking updates. Finally, by increasing the available data rate set further improvements in energy savings can be attained.

Acknowledgments

The work reported in this paper has formed part of the Green Radio Core 5 Research Programme of the Virtual Centre of Excellence in Mobile & Personal Communications, Mobile VCE, `www.mobilevce.com`. This research has been funded by EPSRC and by the Industrial Companies who are Members of Mobile VCE.

References

[1] G. Fettweis and E. Zimmermann, ICT Energy Consumption - Trends and Challenges, *International Symposium on Wireless Personal Multimedia Communications*, Sept. 2008.
[2] P. Grant and S. Fletcher, Mobile basestations: Reducing energy, *E&T Engineering and Technology Magazine*, Vol: 6, No: 2, Feb. 2011.
[3] S. Vadgama, Trends in Green Wireless Access, *Fujitsu Sci. Tech. J.*, Vol: 45, No: 4, Oct. 2009, Page(s):404 - 408.
[4] K. Pentikousis, In Search of Energy-Efficient Mobile Networking, *IEEE Communications Magazine*, Vol. 48, No. 1, Jan. 2010, Page(s):95 - 103.
[5] M.A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, Optimal Energy Savings in Cellular Access Networks, *IEEE International Conference on Communications Workshops, 2009*, Jun. 2009, Page(s):1 - 5.
[6] M. A. Marsan and M. Meo, Energy efficient wireless Internet access with cooperative cellular networks, *Computer Networks*, Vol. 55, No. 2, Feb. 2011, Page(s):386 - 398.
[7] R. Pabst, et al, Relay-Based Deployment Concepts for Wireless and Mobile Broadband Radio, *IEEE Communications Magazine*, Vol. 42, No. 9, Sept. 2004, Page(s):80 - 89.
[8] L.M. Correia, et al, Challenges and enabling technologies for energy aware mobile radio networks, *IEEE Communications Magazine*, Vol. 48, No. 11, Nov. 2010, Page(s):66 - 72.
[9] C. Comaniciu, B.M. Narayan, H. Vincent Poor, and J.M. Gorce, An Auctioning Mechanism for Green Radio, *Journal of Communications and Networks*, Vol. 12, No. 2, April 2010, Page(s):114 - 121.
[10] P. Kolios, V. Friderikos and K.Papadaki, Future Wireless Mobile Networks: Energy Consumption and Resource Utilization in Mechanical Relaying, *IEEE Vehicular Technology Magazine*, Vol. 6, No. 1, Mar. 2011, Page(s):24 - 30.
[11] P. Kolios, V. Friderikos and K.Papadaki, Inter-Cell Interference Reduction via Store Carry and Forward Relaying, *IEEE Vehicular Technology Conference Fall*, Sept. 2010.
[12] P. Kolios, V. Friderikos and K.Papadaki, Load Balancing via Store-Carry and Forward Relaying in Cellular Networks, *IEEE Global Telecommunications Conference*, Dec. 2010.
[13] Wu Hongyi, Q. Chunming, S. De and O. Tonguz, An Integrated Cellular and Adhoc Relaying System : iCAR, *IEEE Journal on Selected Areas in Communication, SI on Mobility and Resource Management in Next Generation Wireless System*, Vol. 19, No. 10, Oct. 2001, Page(s):2105 - 2115.
[14] 3GPP, Technical Specification Group Radio Access Network Study on Minimization of drive-tests in Next Generation Networks, *3GPP*, TR 36.805 v9.0.0, Dec. 2009.
[15] J. Nelson and M. Gupta, An EM technique for multiple transmitter localization, *IEEE Conference on Information Sciences and Systems (CISS07)*, Mar. 2007, Page(s):610 615.
[16] I. Lequerica, P.M. Ruiz and V. Cabrera, Improvement of vehicular communications by using 3G capabilities to disseminate control information, *IEEE Network*, Vol. 24, No. 1, Jan. 2010, Page(s):32 - 38.
[17] D. Hay and Giaccone, P., Optimal Routing and Scheduling for Deterministic Delay Tolerant Networks, *International conference on Wireless On-Demand Network Systems and Services*, Feb. 2009, Page(s):25 - 32.
[18] S. Merugu, M. Ammar and E. Zegura, Routing in space and time in networks with predictable mobility, *Georgia Institute of Technology, GA*, Tech. Rep. GIT-CC-04-07, 2004.
[19] H. Holma and A. Toskala, LTE for UMTS OFDMA and SC-FDMA Based Radio Access, *John Wiley & Sons, Ltd*, 1st. Ed., 2009, Page(s):341.
[20] S. Peters and R. Heath, The future of WiMAX: Multihop relaying with IEEE 802.16j, *IEEE Communications Magazine*, Vol. 47, No. 1, Jan. 2009, Page(s):104 - 111.
[21] S. Parkvall, et al., LTE-Advanced - Evolving LTE towards IMT-Advanced, *IEEE Vehicular Technology Conference*, Sept. 2008, Page(s):1 - 5.

[22] Cui. Shuguang, A. Goldsmith and A. Bahai, Energy-constrained modulation optimization, *IEEE Transactions on Wireless Communications*, Vol. 4, Issue 5, Sept. 2005, Page(s):2349 - 2360.
[23] GW. Miao, et al., Cross-layer optimization for energy-efficient wireless communications: a survey, *Wireless Communications & Mobile Computing*, Vol. 9, No. 4, Apr. 2009, Page(s):529 - 542.
[24] R. Mangharam, et al., Optimal fixed and scalable energy management for wireless networks, *IEEE International Conference on Computer Communications*, Vol. 1, Mar. 2005, Page(s):114 - 125.
[25] S. Chakraborty, Y. Dong, D. Yau and J. Lui, On the Effectiveness of Movement Prediction to Reduce Energy Consumption in Wireless Communication, *IEEE Transactions on Mobile Computing*, Vol. 5, No. 2, Feb. 2006, Page(s):157 - 169.
[26] Y. Dong, WK. Hon, D. Yau and JC. Chin, Distance Reduction in Mobile Wireless Communication: Lower Bound Analysis and Practical Attainment, *IEEE Transactions on Mobile Computing*, Vol. 8, No. 2, Feb. 2009, Page(s):276 - 287.
[27] A. Venkateswaran, V. Sarangan, T. La Porta and R. Acharya, A Mobility-Prediction-Based Relay Deployment Framework for Conserving Power in MANETs , *IEEE Transactions on Mobile Computing*, Vol. 8, No. 6, Jun. 2009, Page(s):750 - 765.
[28] K. Fall, A Delay-Tolerant Network Architecture for Challenged Internets, *Intel Research*, Technical Report IRB-TR-03-003, Feb 2003.
[29] K. Fall and S. Farrell, DTN: An Architectural Retrospective, *IEEE Journal on Selected Areas in Communications*, Vol. 26, No. 5, Jun. 2008, Page(s):828 - 836.
[30] C. Chen and Z. Chen, Exploiting Contact Spatial Dependency for Opportunistic Message Forwarding,*IEEE Transactions on Mobile Computing*, Vol. 8, No. 10, Oct. 2009, Page(s):1397 - 1411.
[31] J. Zhao and G. Cao, VADD: Vehicle-Assisted Data Delivery in Vehicular Ad Hoc Networks,*IEEE International Conference on Computer Communications*, Apr. 2006, Page(s):1 - 12.
[32] W. Zhao, et. al, Capacity Enhancement using Throwboxes in DTNs, *IEEE International Conference on Mobile Ad hoc and Sensor Systems*, Oct. 2006, Page(s):31 - 40.
[33] N. Banerjee, M.D. Corner and B.N. Levine, An Energy-Efficient Architecture for DTN Throwboxes, *IEEE International Conference on Computer Communications*, May 2007, Page(s):776 - 784.
[34] M-R. Ra, J. Paek, A.B. Sharma, R. Govindan, M.H. Krieger and M.J. Neely, Energy-Delay Tradeoffs in Smartphone Applications, *ACM MobiSys*, June 2010.
[35] X. Zhuo, W. Gao, G. Cao, Y. Dai, Win-Coupon: an Incentive Framework for Cellular Traffic Offloading, *IEEE International Conference on Network Protocols (ICNP)*, 2011.
[36] B. Han, Pan Hui, V.S. Anil Kumar, M.V. Marathe, G. Pei and A. Srinivasan, Cellular Traffic Offloading through Opportunistic Communications: A Case Study, *ACM CHANTS*, Sept. 2010.
[37] A. Chaintreau, A. Mtibaa, L. Massoulie and C. Diot, The Diameter of Opportunistic Mobile Networks, *ACM CoNEXT*, Dec. 2007.
[38] J. Reich, A. Chaintreau, The Age of Impatience: Optimal Replication Schemes for Opportunistic Networks, *ACM CoNEXT*, Dec. 2009.
[39] T. Kosch, et al., Communication architecture for cooperative systems in Europe, *IEEE Communications Magazine*, Vol. 47, No. 5, May 2009, Page(s):116 - 125.
[40] E. Schoch, F. Kargl, M. Weber and T. Leinmuller, Communication patterns in VANETs, *IEEE Communications Magazine*, Vol. 46, No. 11, Nov. 2008, Page(s):119 - 125.
[41] M. Gerla and L. Kleinrock, Vehicular networks and the future of the mobile internet, *Computer Networks*, Vol. 55, No. 2, Feb. 2011, Page(s):457 - 469.
[42] R.A. Uzctegui and G. Acosta-Marum, Wave: a tutorial, *IEEE Communications Magazine*, Vol. 47, No. 5, May 2009, Page(s):126 - 133.
[43] H. Mustafa and Y. Zhang, Vehicular Networks: Techniques, Standards, and Applications, *CRC Press*, Apr. 2009.
[44] C.X. Wang, X. Cheng and D. Laurenson, Vehicle-to-vehicle channel modeling and measurements: recent advances and future challenges, *IEEE Communications Magazine*, Vol. 47, No. 11, Nov. 2009, Page(s):96 - 103.
[45] S. Wang, J Min and B. Yi, Location Based Services for Mobiles: Technologies and Standards, *IEEE International Conference on Communications*, Tutorial, May 2008.
[46] Google, The Mobile Movement: Understanding Smartphone Users, *Google/IPSOS OTX MediaCT*, Apr. 2011, www.google.com/think/insights.
[47] C. Botezatu and C. Barca, Intelligent vehicle safety-eCALL, *e-Society*, 2008.
[48] T. Jansen, et al. Handover parameter optimization in LTE self-organizing networks, *IEEE Vehicular Technology Conference*, Sept. 2009.
[49] Y. Zhao, Standardization of Mobile Phone Positioning for 3G Systems, *IEEE Communications Magazine*, Vol. 40, No. 7, Jul 2002, Page(s):108 - 116.
[50] ETSI 3rd Generation Partnership Project, LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Stage 2 functional specification of user equipment positioning in E-UTRAN, *ETSI Technical Specification 136 305 V9.0.0 Release 9*, Oct. 2009.
[51] 3GPP, Technical Specification Group Radio Access Network Study on Minimization of drive-tests in Next Generation Networks, *3GPP*, TR 36.805 v9.0.0, Dec. 2009.
[52] K. Jooyoung, et al, A novel location finding system for 3GPP LTE, *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept. 2009, Page(s):3213 - 3217.
[53] I. Constandache, R.R. Choudhury and I. Rhee, Towards Mobile Phone Localization without War-Driving, *IEEE INFOCOM*, Mar. 2010.
[54] A. Thiagarajan, L.S. Ravindranath, H. Balakrishnan, S. Madden and Lewis Girod, Accurate, Low-Energy Trajectory Mapping for Mobile Devices, *USENIX Symp. on Networked Systems Design and Implementation*, Mar. 2011.
[55] A. Schrijver, Theory of Linear and Integer Programming, *Wiley Press*, April 1998.

Panayiotis Kolios is a member of the Institute of Telecommunications at King's College London. The reported work has been part of his studies while working towards the completion of his PhD degree in Telecommunications Research. His research interests revolve around the broad scope of wireless networking, Internet routing, multimedia communication and network science. He is an active member of IEEE, contributing to a number of technical and professional activities within the association.

Vasilis Friderikos lectures at the Institute of Telecommunications at King's College London in the UK. His research interests lie broadly within the closely overlapped areas of wireless networking, mobile computing, and architectural aspects of the Future Internet. He has published more than 100 papers in the above scope research areas and actively involved in flagship IEEE conferences such as Globecom, ICC and PIMRC.

Katerina Papadaki is a Lecturer in the Management Science Group, Department of Management at the London School of Economics. She received her PhD from Princeton University from the Department of Operations Research and Financial Engineering. A major component of her research has been in developing algorithms to solve discrete deterministic and stochastic optimization problems that arise in dynamic resource allocation problems with applications involving physical resources (transportation networks), and radio resource allocation (wireless communication networks). Recently, her research attention has been on game theoretic problems. These include fair allocation of resources in telecommunication networks and network patrolling games with applications in detecting terrorist activities. She is associate editor of Optimization Letters and a member of INFORMS and IEEE.