IEEE CAS Workshop on Wireless Communications and Networking, September 5-6th, 2002, Pasadena, CA

COST METRICS FOR DECISION PROBLEMS IN WIRELESS AD HOC NETWORKING

Michael Angermann and Jens Kammann

German Aerospace Center (DLR), Institute of Communications and Navigation
P.O. Box 1116, D-82230 Wessling, Germany
{Jens.Kammann, Michael.Angermann}@dlr.de, Fax: + 49 8153 281871

ABSTRACT

The main contributions of the presented paper are an in-depth discussion of the special conditions existing for multiple wireless connection scenarios. We evaluate reasonable structures for cost functions and optimized decision processes and present a simulation to demonstrate the influence of the various parameters. Finally we discuss controller strategies that implement these cost functions.

1. INTRODUCTION

Current and future mobile and portable devices such as smartphones or laptops are showing a clear tendency to incorporate not just one, but a multitude of wireless connection technologies per device. Many current phones have the ability to choose between standard circuit-switched data (CSD), high speed circuit switched data (HSCSD), general packet radio service (GPRS) and Bluetooth to transfer data. On Laptops, wireless LAN is often an additional alternative. If these devices are to connect to various networks in an ad hoc fashion to automatically perform tasks, efficient decision processes are needed to choose among these possibilities to optimize network usage and network availability both from the user's and the system's perspective. In conjunction with handover procedures[1][2], pro-active retrieval of data (prefetching)[3] has the potential to reduce overall cost under frequently changing network conditions, despite the fact of an overall increased network load. Channels, timing and issued requests have to be carefully decided upon. We treat these decision processes as an optimization problem, for which suitable cost functions are needed. Typical metrics employed in fixed network scenarios, such as hop count are not well applicable in many wireless scenarios. Parameters like power consumption, call setup times, significant latency due to sophisticated channel coding methods and others are rarely addressed in previous network optimization research. In this paper we point out the requirements for useful metrics and propose sample cost-functions of system parameters that result in such metrics. In contrast to many research contributions that assume the optimization process to be situated in the network with overall knowledge of network state and traffic demand we investigate a decentralized optimization from the individual client perspective. We propose this as a reasonable approach for its technical feasibility, as it can be implemented in software in or on top of the mobile device's operating system. We will present results derived from simulations at bit-, packet- and request/response level. An insight determined from these results is parameter sensitivity of the cost functions. The simulations are performed with dependent traffic sources for uplink and downlink in order to achieve meaningful results for the performance under typical request/response oriented protocols such as HTTP.

This work was done within the Heywow Project. More Information can be found at http://www.heywow.com

2. SYSTEM MODELING

As shown in figure 1 our proposed model is divided in three parts: Traffic Model G_T, Network Model G_N and Cost Model G_C.

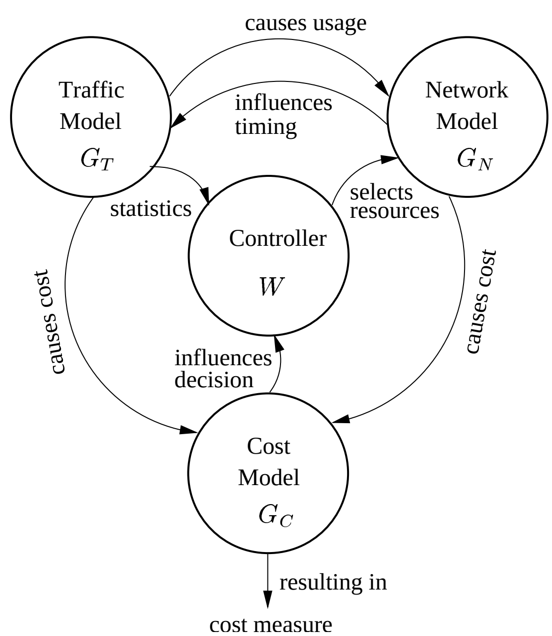

Figure 1. Model overview

2.1. Traffic Model

Our traffic model assumes HTTP traffic: Each request r_{req} is followed by a response r_{res} and a viewing time T_{View} The distribution of request and response sizes as well as the viewing time are based on empirical results derived by [4]. Figure 2 shows detailed timing of one request with its associated response: T_δ denotes the network delay, T_τ is time it takes for transferring the data, depending on actual capacity and queue length for the given channel. T_S is the time the server need for processing the request and preparing the response data. The response follows the same nomenclature. After the viewing time T_{View} has expired, the next request may be issued. The simulation also considers primary and secondary requests which are found when browsing the Web: The primary request usually fetches the main HTML page, whereas several secondary requests may be necessary to download all embedded objects such as images and scripting parts.

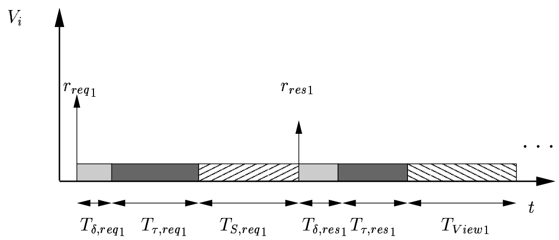

Figure 2. Traffic Model (1)

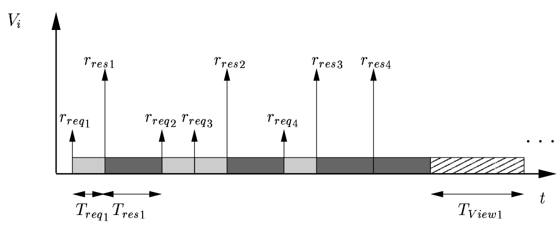

Figure 3. Traffic Model (2)

2.2. Network Model

The multitude of connection types available on a mobile device e.g. Bluetooth, WLAN, GSM and others are modeled by N independent *channels*. To accurately represent the case where the availability of these channels show temporal fluctuations we model each individual channel by a two-state first order, homogeneous and stationary Markov model.
Their state $S_i, i \in \{1, 2\}$ represents an *available* "1" or *unavailable* "0" channel.

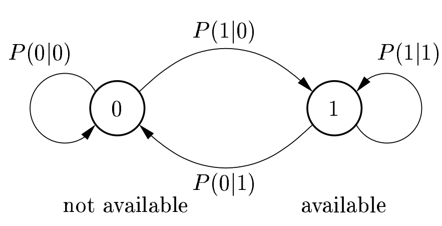

Figure 4. two state Markov channel model

This model randomly changes its state at discrete times $k \cdot T_{s,channel}$. Its behavior is completely defined by the channel sampling interval $T_{s,channel}$ and a pair of state transition probabilities $P(S|0), P(S|1)$.

To illustrate how this Markov model's parameters are determined we present a small example:
We choose the sampling time of the channel state $T_{s,channel} = 0.5$ s; the mean duration in the available state shall be $\overline{T}_{d,1} = 10$ s and the network shall be unavailable with probability $P(0) = 0.25$
The probability mass function of the duration d in state S_i is (with $p_{i,j} = P(j|i)$)

$$
\begin{aligned}
f_d(d) &= P(\{\overbrace{S_i, S_i, \ldots, S_i}^{d \text{ times}}, S_j \neq S_i\}|S_i) \\
&= p_{i,i}^{d-1} \cdot (1 - p_{i,i}).
\end{aligned}
\tag{1}
$$

From this we calculate the *mean* duration in state S_i

$$
\begin{aligned}
\overline{d}_i &= \sum_{d=1}^{\infty} d \cdot f_d(d) \\
&= \sum_{d=1}^{\infty} d \cdot p_{i,i}^{d-1} \cdot (1 - p_{i,i}) = \frac{1}{1 - p_{i,i}}.
\end{aligned}
\tag{2}
$$

From eq.2 and $\overline{T}_d = T_s \cdot \overline{d}_i$ follows

$$
P(1|1) = 1 - \frac{1}{\frac{\overline{T}_{d,1}}{T_s}} = 1 - \frac{1}{10\text{s}/0.5\text{s}} = 0.95 \tag{3}
$$

Assuming stationarity we may use

$$
P(0) = P(0) \cdot P(0|0) + (1 - P(0)) \cdot P(0|1). \tag{4}
$$

From which follows

$$
\begin{aligned}
P(0|0) &= \frac{P(0) - (1 - P(0))(1 - P(1|1))}{P(0)} \\
&= \frac{0.25 - 0.75 \cdot 0.05}{0.25} = 0.85
\end{aligned}
\tag{5}
$$

All necessary parameters of the Markov model are now determined. A slice of a simulated temporal availability of a sample Bluetooth channel with parameters chosen according to the example is depicted in figure 5 in combination with a second (GSM) channel with $\overline{T}_{d,1} = 600$ s and $P(0) = 0.05$.

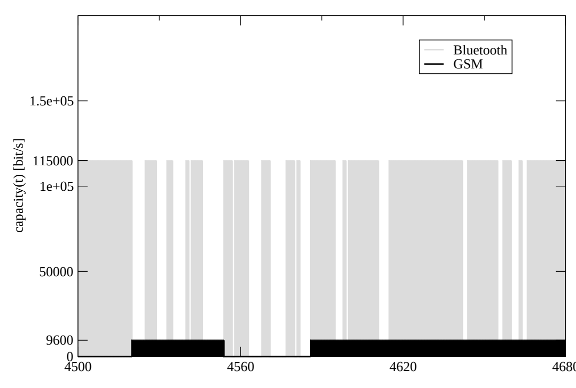

Figure 5. Channel Availability

We decided to decouple the sampling interval $T_{s,channel}$ of the channel state from the actual simulation sampling interval T_s for two reasons: a) the possible duration within a state is lower-bounded by $T_{s,channel}$. Thus, by choosing $T_{s,channel}$ suitably large, we can suppress unrealistic short "spikes" in the channel availability. b) Channel availability traces generated from more complex e.g. geometric models [5][6] or actual field measurements can be used within the simulation without having to lower the simulation sampling rate to their usually low rates.

2.3. Cost Model

Our proposed cost model considers the cost for data transfer (network cost), cost for user waiting time and energy consumption:

$$\begin{aligned} c(i) &= \underbrace{\alpha_V(i)V + \alpha_T(i)T(i)}_{\textit{network}} + \underbrace{\beta T(i)}_{\textit{user}} \\ &\quad + \underbrace{\gamma_V(i)V + \gamma_T(i)T(i)}_{\textit{energy}} \end{aligned} \tag{6}$$

with

$$T(i) = T_\delta(i) + T_\tau(i) \tag{7}$$

and

$$T_\tau(i) = \frac{q(i)}{\mathrm{E}\{C(i)\}} \tag{8}$$

where $q(i)$ denotes the actual queue-length of channel i and $\mathrm{E}\{C(i)\}$ is the expectancy of available capacity for ch.i.

2.3.1. *Network cost (α)*

Circuit switched channel usually are billed by usage time and number of channels used. The time may be quantified from one second to minutes, with the first interval being longer than subsequent ones (e.g. 60/1 means first minute is charged in full, afterwards the the price per minute is charged in increments of seconds). The current model assumes a link ready for payload right after billing of air-time starts which is the case if the called destination uses ISDN. There might be a one to two second delay for negotiating the layer 3 protocol (e.g. V.110 bit rate adaptation). The called destination is an analog modem, air-time applies already for the time the mobile phone networks and the destination negotiate the carrier. This time can be reduced by disabling the auto-negotiating mode at the destination but still remains 10 to 30 seconds before payload data transfer can start. Packet switched communication is usually billed by volume of transferred data. The current model takes into account the packet size during a session but we assume that a mobile phone if it switches from packet to circuit switched mode it will park the packet switched session and is able to resume the same session afterwards again. If the devices disconnects the packet session first, another quantization takes place (e.g. 10 or 100 KByte increments) and re-establishing the session takes several seconds. We assume the scope of the model is only a fraction of a billing period, so we do not consider any included air-time or volume contingent

2.3.2. *Cost for user time (β)*

Research has shown that a human user is only willing to wait a certain time for a response of a service. In the case of browsing the web using a fixed Internet connection there is a threshold of around 7 seconds within a user expects a response. In mobile networks this time is longer, but response times over 20 seconds generally cause the user to think a connection might be broken. If the "user" is an agent application that automatically fetches information, the maximum acceptable time is the time which allows the application to perform a given task in a timely manner (e.g. an E-Mail application is less critical than a stock ticker or a personal travel assistant). Current simulation considers only the linear term $\beta T(i)$, but can easily extended to a series expansion $(\beta_0 + (\beta_1 - t_1) + (\beta_2 - t_2)^2 + \ldots + (\beta_n - t_n)^n)T(i)$ to approximate the non linearity.

System	idle	active
Circuit Switched	medium	high
Packet Switched	medium	high
W-LAN	medium	medium
Bluetooth	low	low

Table 1. Power Consumption

2.3.3. *Energy cost (γ)*

Limited battery power of mobile devices requires deliberated selection of the communication channel. Table 1 gives a qualitative overview of the power consumption of common wireless networks. We do not distinguish between the communication related and the computation related energy consumption [7] because we do not attempt to modify existing wireless standards, i.e .by improving MAC layer protocols. A connection is "active" if a link is established and data is being transferred (corresponding to the parameter T_τ of the traffic model". An idle connection also contributes towards the total cost and is considered by the parameter T_{View} of the traffic model.

3. SIMULATIONS

Formally we try to solve a optimization problem by finding an optimal controller W which minimizes overall system cost. Given are cost model $G_{T,N,C}$ and the quality function

$$J = \mathbf{c}^T \cdot \mathbf{x} \tag{9}$$

where c denotes a vector of length N with N representing the number of available channels. Each element $c(i)$ of this vector is described by eq.(6).
For our simulation we assume that each request or response is sent without further division using exactly one channel:

$$\mathbf{x} = (x_1, x_2, \ldots, x_i, \ldots, x_N)^T, \quad x_i = 1, x_j = 0,\ \forall\, j \neq i \tag{10}$$

One simple minimization of cost is to select the channel with the lowest current cost:

$$i = arg\ min\, c(i) \tag{11}$$

Figure 6 shows the simulation snapshot of two sample webpage requests. We assume two wireless channels, Bluetooth and circuit switched GSM with their availability given by the network model of chapter 2.2. The first webpage request of the given snapshot occurs at $t = 92,5s$ and consists of one primary and one secondary request and its corresponding responses. The second one occurs at $t = 93,8s$ and consists of three secondary requests/responses along

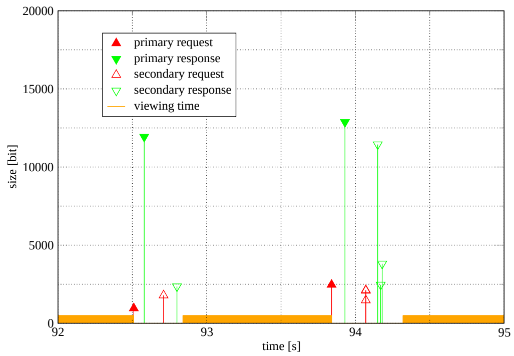

Figure 6. Snapshot of sample HTTP requests and responses

work, user and energy according to the cost model proposed in chapter 2.3 as the outcome of the simulation. The total cost for multiple simulation runs with varying model parameters led to some preliminary controller strategies presented in the following chapter.

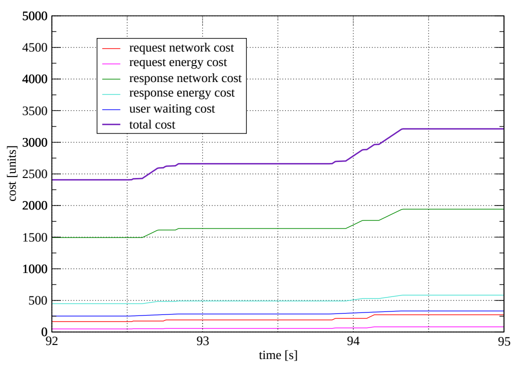

Figure 7. Total cost and individual contributions

4. CONTROLLER STRATEGY

The simulations described in the previous chapter assumed a simple controller W which decides on a request/response level based on the current cost for each channel and the time a request occurs. This is a reasonable trade off between complexity (the controller runs usually on mobile devices) and efficiency. The global minimum could only be found if all request and their time of occurrence would be known in advance. More realistically, the controller could learn the statistics of the requests over time. This is likely to further reduce cost, especially if request originate from software agents[8] whose behavior is easier to predict than a human being surfing the web. The same is true for improved statistics for the channel availability: The expected value of the capacity is easy to estimate, but leaves room for improvement, e.g. large requests should not be send over a highly volatile channel. Also each cost term should have an upper threshold. If a request exceeds this limit, it should be considered as a failed request (e.g. execution time too long or too expensive). This allows to balance the cost parameters under an application dependent acceptable loss ratio.

5. SUMMARY AND FURTHER WORK

Verifying the applicability of the proposed metrics and methods, as well as calibrating simulation parameters against measurements are the next important steps in our work. The ultimate goal is to have an mobile device which automatically configures each available wireless channel. The user should have the possibility to affect the controller's behavior via a simple user interface as shown in figure 8.

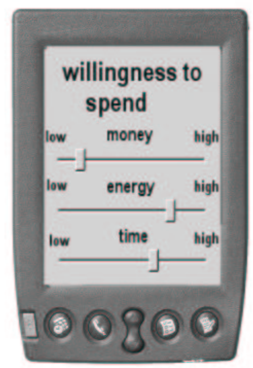

Figure 8. Mock-up User Interface

6. REFERENCES

[1] J. Kammann and T. Blachnitzky, "Split-proxy concept for application layer handover in mobile communication systems," in *Proceedings of the 4th IEEE Conference on Mobile and Wireless Communications Networks (MWCN 2002), Stockholm, Sept.*, 2002.

[2] M. Stemm and R. H. Katz, "Vertical handoffs in wireless overlay networks," *Mobile Networks and Applications*, vol. 3, no. 4, pp. 335–350, 1998.

[3] M. Angermann, "Analysis of speculative prefetching," *ACM Mobile Computing and Communications Review*, vol. 6, April 2002.

[4] B. A. Mah, "An empirical model of HTTP network traffic," in *INFOCOM (2)*, pp. 592–600, 1997.

[5] I. Akyildiz and W. Wang, "A dynamic location management scheme for next-generation multitier pcs systems," in *IEEE Transactions on Wireless Communications, Vol.1, No.1*, January 2002.

[6] P. Escalle, V. Giner, and J. Oltra, "Reducing location update and paging costs in a pcs network," in *IEEE Transactions on Wireless Communications, Vol.1, No.1*, January 2002.

[7] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J.-C. Chen, "A survey of energy efficient network protocols for wireless networks," *Wireless Networks*, vol. 7, no. 4, pp. 343–358, 2001.

[8] T. Strang and M. Meyer, "Agent-environment for small mobile devices," in *Proceedings of the 9th HP OpenView University Workshop (HPOVUA)*, HP, June 2002.