Paper Analyses

Paper: "Application-Driven Cross-Layer Optimization for Video Stream Over Wireless Network"

Main idea: Optimally allocate resources and adapt to dynamically changing environments for mobile multimedia applications in order to maximize user satisfaction.

Previous work has mainly focused on single layers, either in a top-down approach (typically passes priority labels to lower layers which performs class-based queuing and priority-based transmission) or a bottom-up approach (typically exploits information about the current channel situation to adapt the transmission policy of the application). However, some of the ongoing research focus on joint optimization of the physical and data link layer. CLO including the application layer had just recently appeared.

By making effective abstractions of the application, data link and physical layers, they jointly optimize parameters of multiple layers. Their CLA has N layers and a cross-layer optimizer (CLO) which uses an application-based objective function.

First an abstraction of layer-specific parameters is computed. This is the first step (layer abstraction). Then the optimizing itself starts. The proposed CLO finds the values of layer parameters that optimize a specific objective function (in this case, it is a function of the expected video reconstruction quality of multiple users). All that is left, is the reconfiguration. The optimal values are distributed to the corresponding layer. These three steps complete the CLO.

They classify the parameters into four groups based on whether the parameters are tunable, and whether they are descriptive or abstracted. DT (directly tunable) parameters can be set directly, while IT (indirectly tunable) parameters may be changed as a result of the setting of DT parameters. Non-tunable parameters that can be read by the CLO are descriptive parameters. Abstractions of these three, are the abstracted parameters.

A group of pictures (GOP) is a sequence of groups of consecutive frames. A video stream is usually encoded like this. There are three different frames: I-frame (encoded independent of other frames), P-frames and B-frames (last two are differentially encoded with respect to other frames in the same GOP. I-frames are the most important in the way that if they are lost, the distortion will be higher than by losing a B- or P-frame. The whole idea with different frames are pretty clever. As figure 2 states: with 16 frames,1 I-frame followed by 15 P-frames, the loss of the I-frame will cause a high reduction in quality. On the other hand, the reduction becomes lower and lower for each packet that is correctly received. By this, we see that loss of a particular frame in one video sequence may have little influence on the reconstruction value, while it in another video sequence can cause a high reduction. This is one the key problems that the authors address: They use dynamic resource allocation across multiple users.

Even though the channel conditions change, timely delivery of video is expected, even in wireless video streaming which is one of the big challenges. In order to optimize the end-to-end quality in mobile networks and efficiently use the network resources, we need adaptability to change application and network characteristics on all layers. The application layer can adapt to varying network characteristics by adequate processing. Dynamic rate adaption at the server is one example.

The authors tell the reader to look up a paper in order to get an overview of properties and challenges that come with video streaming. This seems like easy way out, but the reader is supposed to have some background, and the paper cannot explain everything. If it had done so, the paper could be a lot longer. The length is good, but making it longer would not necessarily be a good thing if they only added rather basic explanations. However, they could probably have mentioned some high-level pointers in a couple of sentences to fresh up the memory.

The so-called peak signal-to-noise ratio (PSNR) that closely represents user-perceived video quality. And by this fact, they use this to define an objective function for optimization of video streaming delivery systems.

At the radio link layer, four key parameters are abstracted: transmission data rate, transmission packet error rate, data packet size, and channel coherence time. These make up the abstracted tuple that describes the radio link layer. The application layer abstraction is used so that the optimizer can be aware of the effects of lower-layer parameters on the application layer. "Rate distortion profile" is a convenient abstraction tool.

It has been shown that a cross-layer architecture have a relevant performance gain over the traditionally layered one. This paper introduced an application-oriented cross-layer optimization concept for wireless video streaming. Information exchange was performed in both directions (that is, top-down and bottom-up). The process of parameter abstraction is one of the most important ones. This process' contribution is that it helps keeping the number of parameter tuples to be evaluated reasonable. As they say themselves, the definition of the optimization metrics in CLD is an important issue. And the metrics of different types of media are hard to compare. CLO improves network capacity and increases the number of users served. But, it is hard to fully guarantee satisfaction of QoS constraints in all scenarios with a distributed and highly dynamic environment.

The paper was not that straightforward. It has some complex information. In their simulation, they had two pre-encoded videos. I did not see anything about the cost of this in the "Cost Analysis"-section. If all videos are to be pre-encoded, a server with a lot of videos would probably have storage problems. And if they have to be encoded on-the-fly, it would require more processing-power. This should have been in the analysis.

Abbreviations (shortened)

B3G – beyond-third generation
CLA – cross-layer architecture
CLD – cross-layer design
CLO – cross-layer optimizer
GOP – group of pictures
OFDM – orthogonal frequency-division multiplexing
PSNR – peak signal-to-noise ratio
QoS – quality of service
TDMA – time-division multiple access