CS 641 Takehome Midterm Exam

Ahmet Burak CAN
abc@purdue.edu

Question A

Paper: An Evolution of Transmitting Compressed Images In a Wide Are Network
Melliyal Annamalai, Bharat Bhargava

1. What is the reseach question?

Distributed applications that deal with large multimedia objects spend much time to transmit data over network. Thus, reducing the transmission costs improve response time and system efficiency. One way to improve response time is reducing data size. Lossy compression techniques for multimedia data can decrease data size in charge of losing some information. Multimedia data is more tolerant to lose some parts of data. In this paper, impact of reducing image quality, in other words reducing filesize, to achieve better transmission times over wide area network is studied. To do it, effects of some parameters on transmission time should be identified, such as, transmission distance, image attributes, time of day, computation power, physical media, storage, user defined quality level, response time. Relationship between these parameters and transmission time should be found. Transmission time of an image basically consists of three main processes: transmission over network, accessing to requested image data on sender (getting image file, compressing it etc.) and displaying image on receiver. These parts should be studied separately and their effects to each other should be found.

2. Specify ideas that can make this area succeed commercially.

The concepts explained in this paper can also be expanded for video and audio data transmission but it needs further research. I will explain some ideas that can make important effects on transmission costs for multimedia data delivery, not only restricted with image delivery. I will look at the benefits of reducing costs of transmission of multimedia data.

Every enterprise organization use database to keep data of previous organizational works. Several decades ago, these data were consists of text and numeric data related with some classified works. However, today, this kind of data is not sufficient to satisfy needs of future works, keeping multimedia data is gaining importance everyday. Such as, NASA collects some statistical data about space observations everyday and uses this information for future experiments. Besides this, NASA also collects terabytes of image data about space observations. Comparing text and numeric data retrieval, retrieval of image data takes much time. Reducing retrieval time of an image data makes big help in scientific research.

Internet carries multimedia data increasingly everyday, some times carried multimedia data is more than text data. Multimedia related companies that deliver multimedia data over Internet will become popular in near future. Music companies are working on legal delivery of songs over Internet. In future, movie companies may setup movie saloons over Internet which enables to watch movies on Internet. It is already possible, but, as I know, there is no commercial and legal application now. Gnutella like peer-to-peer networks already enable sharing of movies by individuals, but protection of intellectual property issues should be solved. Some TV companies and radio stations started to deliver some of their programs. Maybe, in ten years, they will give their whole programs from Internet.

One of the important emerging concepts is Digital Libraries. Digital libraries provide online access to text and multimedia data in integrated manner. Thus, they are one of the major users of Internet infrastructure. For now, IEEE and ACM and some other publishers make available their old printings from Internet. These are some important steps for future Digital Libraries, they lack to provide any kind of information available, their content are limited with publisher's own printings. Actually WWW can be considered as a Digital Library, it's content can be searched with search engines and content include almost every kind of data. However, the results returned from search engines are generally dependent to simple text searches over the web pages around the world. Some content extraction from multimedia data should be done to be able consider these applications as fully functional with multimedia data. At this point, pattern recognition and data mining research gain importance.

Related to Digital Libraries, Image and Video databases, Geographical Information Systems are big users of multimedia data. These applications need to deal vast amount of multimedia data, hence transmission costs are crucial for their system efficiency.

Almost all of the entertainment companies need to use multimedia data to attract more people. Such as, travel companies need to advertise their tours and Internet is one of the most important environments for this.

Distance education is also developing concept in universities and companies. Instead of carrying people to educate, bringing education material to them is much cheaper and comfortable. There are some universities that have started to deliver their course contents for distance education. Companies have to provide some job related education to compete with others. Instead of completely leaving weeks of employees to education, providing some online content and letting them to study on their spare time is much cost effective for companies. It also decreases the travel and vacation costs of educations.

Video conferencing is another application that uses network bandwith to deliver video data. Companies, scientists, especially in medicine, for example in a surgery operation, use video conferencing to exchange information.

3. Give results/experiments that are based on one of the ideas.

The experiments are done on two image formats: GIF and JPEG. Firstly these image formats are studied. GIF is designed as a transmission format for images and it's a lossless compression technique. JPEG uses a lossy compression which means a compressed image can not be reconstructed from JPEG format to exact original image. Quality of JPEG image can be adjusted by applying different compression ratios. Some concepts and details about these image formats are also explained in the paper.

A response time formulation for experiments is defined:

$$\textit{response time} = t_a + t_t + t_d$$

Some parameters are defined that may effect this response time. Some of them are under user control, some of them are not. The uncontrollable parameters:

- Transmission Distance: In Internet a sender and receiver may be in the same LAN or may be on very distant parts of the world. Hence, this parameter can take very different values.
- Image Attributes: Image contents may make difference on size and quality of compressed image. Photographs, detailed pictures can be compressed better using JPEG format. However, such as for simple cartoons, line drawings or text containing pictures, GIF format is better. If content of image can be classified with some parameters, suitable compression technique and ratio can be determined.
- Time of Day: This parameter is directly related with network traffic ratio. In week days and in work hours, networks are generally busy, so transmission time will be high. This parameter is not easy to figure out, but it can be determined with samplings of Internet traffic.
- Computation Power: Compression, decompression and display of images on sender and receiver take some time. Compressing images reduce transmission time, but if it takes excessive time and does not give much advantage in total response time, it may be better not to compress an image.
- Physical Media: The transmission media is also important for determining transmission time on network. This parameter determines bandwidth, forwarding delay, packet drop rates etc.
- Storage: If storage is not a constraint, most frequently requested quality levels of images can be stored together to reduce load time of image at sender site.

Besides this uncontrollable user parameter, two user controllable parameters are defined:

- Quality level: Different kind of users may need different qualities of images. Thus, assuming an average quality level for every user is not a proper approach, some users may need quality images, some may not be tolerant for delay and moderate quality may be enough for them. Frequently requested quality levels give hints to store which quality level of image should be stored.
- Response time: Like quality level, response time can be a user parameter. These two parameters are related reversely proportional.

As mentioned above, JPEG does not perform well for every images, it depends on the content of the image. Number of colors, number of sharp edges and amount of different colors etc. all effect compressed image quality. For some image, %30 compression may result a perfect quality image, but for some other it may produce poor quality image. The quality of image and compression ratio should be rated to be able determine maximum compression ratio for a quality level. For one image perfect quality can be obtained with %30 compression ratio, but for other it may be %10. Four levels are defined for image quality: Perfect, Partial Loss of Information, Complete Loss of Regions, New Regions are introduced.

Basically, three main experiments are explained in this paper. 19 different images are used to represent different kind of images.

First experiment is to measure transmission time over the network. For precise measurements, experiments should be done on Internet. However, it's hard to have different accounts on different administrative domains. Because of this, *TCPecho* program is used to send image to remote hosts. This program uses echo protocol which is available on most of the servers on Internet, so experiments can be tried on different network hosts without setting up any program and system. The experiments are repeated on 7 different remote hosts and 300 tries are done on each host. According to this experiment results, the difference in round trip time between small and large image is not significant in LAN environment for users. However, for hosts on a WAN, the filesize makes great impact on transmission time and reducing file size can be meaningful by using compression techniques. Only beyond a certain network distance, compressing a file achieves a significant saving in response time. Additionally, if file size is less than a certain threshold, the file can be retrieved as it is regardless of the network distance. If the network traffic is considered, this threshold can be lower. Thus, in busy networks, to get a certain response time, much tradeoff in quality may need to be done.

Second experiment was to measure the time to perform conversion from GIF format to JPEG format. This time effects t_a parameter in response time formula. The conversion process is repeated 300 times on each sample image by using *cjpeg* application. The results show that conversion time is not directly related with file size. Some images can take longer compression time comparing bigger sized files. Semantics of an image like number of colors, shades, amount of details etc. may also effect this time. As a result, compression time related with both file size and the semantics of image.

Third experiment was to measure the time taken to decompress and display different image files. The experiments are done using *XV* software and the tests are repeated 300 times for each image. Image formats have varying degrees of complexity and display time varies accordingly, in general GIF takes less time than JPEG. However, like in compression, display time is also related with image semantics. For two same sized and same formatted images, one of them may take much time.

As a result of experiments, transmission time, t_t, is dependent to network distance, time of day (in other words network traffic), image quality (in other words file size). Access time, t_a, is dependent to computation power of the sender machine. Display time, t_d, is dependent to computation power of the receiver host machine.

4. What are the experiments need to be conducted to make progress in the area? What tools are available and can be developed?

Firstly, quality levels for users should be studied more elaborately. It's hard to define scientific quality levels. Good picture for one may be in moderate quality for an artwork person. Additionally, users may not be knowledgeable about what quality levels they want. In this case, default quality level can be defined for new users and then the user's requests can be used to determine appropriate quality level in the future requests.

Image content can cause different file sizes and quality levels for same compression ratio. Users can only specify a quality level for a requested image, but it's hard to find exact, or more close to exact, compression ratio for requested quality level. If image compressed excessively, the quality will be low. Otherwise, quality level will be higher than expected and filesize and response time will high. If image content can defined more precisely, maybe more mathematically, better compression ratios can be defined for desired quality levels.

For JPEG format, progressive transmission is not available. Progressive transmission means that lower quality levels can be used to enhance higher quality levels. JPEG can only replace high quality with low quality one, which causes retransmission of some already transmitted information. If progressive transmission could be achieved, it would be a major contribution for many applications.

Similar experiments for video and audio data formats can be done. Like image data, these kinds of multimedia data constitute a lot of traffic on Internet. Effects of using compression on video and audio data should be studied. Video data takes more time to decompress data than image data. Thus, sender and receiver site computation power should be studied more elaborately.

MPEG video format allows progressive transmission of video data. In this format, video data represented as some layers. At the start, a few layers are transmitted, if the user is not satisfied with current quality level, next layers are transmitted and these layers are used to enhance quality of previous scenes. This method and its effects on transmission time can be studied.

Wireless technology seems to be next decade's communication technology. Wireless networks have more error rates comparing to wired networks and their transmission protocol behave different from wired networks. Also, wireless devices may have some limitations, and computation power can be an important criteria. On wired networks, almost all machines have enough power to decompress a JPEG image in a short time (for a user). However, if a handheld device is the receiver, displaying image may take tens of seconds. And user screen may not be sufficient to display every kind of image. Such limitations and different behavior of transmission protocol and media should be studied on the experiments defined in this paper.

Question B

Paper: Characterization of Video Traffic
Rahul Garg

In this paper, characteristics of real-time video traffic generated by video on demand, video conferencing systems, etc. is studied. Standard video coding algorithms JPEG, MPEG and the video conferencing software NV is used in experiments in order to sample video traffic. Several video contents from movies to a class lecture were analyzed.

Firstly, traffic descriptors for video traffic are defined. Traffic descriptors are used in the connection setup phase to determine whether the network has enough resources to satisfy the QOS requirements of the traffic. The leaky bucket model is used to characterize the bit rate of the traces. A leaky bucket descriptor of parameters (σ, ρ) means that the amount of data carried by the stream in any interval of length I is bounded by $\sigma + \rho I$. Another traffic descriptor, burstiness function is used to characterize the burstiness of the video traffic. The burstiness function of a traffic stream is defined as:

$b(I)$ = Maximum amount of data sent in any interval of length I .

If burstiness curves plot the average bit rate in the interval in which maximum amount of data is sent, slope of the curve is a measure of the burstiness of the traffic.

For observing JPEG compression technique, five different video data are studied. These videos have different kind of contents, different picture qualities and different frame rates. At first glance, visual inspection of the bit rates shows that frame sizes of consecutive frames are highly correlated. Because, the size of a compressed frame is essentially determined by the complexity of the image (information content) and consecutive images are generally very similar in visual appearance, complexity and information contents. For example, in one of traces, a lecture is recorded and the bit rate is not changing so frequently. When the instructor writes text on the board, there is more information in the corresponding image and because of a lower compression, the corresponding bit rate increases. During a scene change, the pictures on different scenes are not similar to each other, so the sizes of the compressed frames are not correlated and bit rate changes suddenly. For the duration of a scene, the bit rate remains close to the peak rate and the worst case average bit rate will be close to the peak rate.

Leaky bucket descriptor would require that the service rate be close to the peak rate for JPEG compressed video traffic. In a lower service rate, the bucket size required will be very large. The large bucket is because of the instantaneous arrival rate remains a little higher than the service rate for relatively long periods of time. The peaks of the bit rate occur during the scenes having the high image complexities. A traffic characterized by a large bucket size can potentially inject a large burst of data in the network.

Characteristics of MPEG algorithm is observed on two different MPEG data. MPEG is an inter-frame coding algorithm that uses spatial and temporal redundancy of the video for the compression. The MPEG frames are classified into 3 types - I-frames, P-frames and B-frames. This three different types of frames generated causes high burstiness when observed over short time scale. The short term burstiness smoothes out when the bit rate is averaged over intervals of appropriate length. This length depends upon the periodicity of I-frames in the stream. In order to achieve random access on the frames, I-frames are periodically inserted in MPEG streams. P-frames and B-frames constitute rest of the video and they occur more frequently. After smoothing, the bit rate starts showing dependency on the contents of the video, because of the periodic presence of I-frames in the stream, and their content is dependent to the complexity of pictures. For test cases, when averaged over an interval of length 1 second, the bit rate becomes smoother.

The service rate for MPEG compressed video traffic varies from peak rate to the eventual average rate. The bucket size doesn't increase even if the service rate is less than the peak bit rate. The reason for this is that the largest frame is I-type and it is always followed by a number of small B-type frames. By the time the frame next to the largest I-frame arrives, the bucket has enough space to accommodate it without necessarily serving at peak rate. In case enough bandwidth is not available, congestion would result into a series of localized packet loses. The parameter bucket size of the leaky bucket can also act as a measure of the buffering needed to smooth the short term burstiness.

Characteristics of NV are observed on three different NV traces. NV traffic is highly dependent to the rate control built-in the software. In NV traffic, most of the packets are spaced at 5 ms, but sometimes this time may be as large as 0.5-2 seconds. The reason is that NV stops sending data if it realizes that it is sending the data at high rate. So when NV sends the data at high rate and following periods will be on inactivity to lower the average bit rate. Some periodic packet spacing variations can be attributed to periodic waking up of other processes on the same host, or other periodic activity in the same LAN segment (e.g. routing updates).

The traffic generated by the NV program is highly bursty in short term due to its on-off nature. The long term behavior of NV traffic is highly predictable. Such as, a plot of average bit rate taken over 10 second interval is nearly constant. Constant bit rate over a long time scale is achieved by changing the frame rate of the video being sent. Thus, the characteristics of the NV traffic are fairly independent of the video which is being sent.

As result of the experiments, The appropriate service rate for JPEG compressed video is close to its peak rate. The service rate for MPEG compressed video is approximately equal to the peak rate of MPEG stream after smoothing. The service rate of NV traffic is slightly more than its target sending rate.

The parameter bucket size depends heavily upon the way the application delivers data to the network. In case of JPEG and MPEG compressed video, we assume that the application delivers data in form of frame at a constant frame rate. Therefore, the bucket sizes obtained are close to the size of the largest frame present in the stream. The paper characterizes the traces using a leaky bucket model and shows a way of choosing appropriate leaky bucket parameters. The leaky bucket parameters for constant quality JPEG and MPEG compressed video depend upon the actual video streams.

Burstiness function characterizes the burstiness of the traffic at various time scales. The impact of burstiness on network congestion is discussed. It is shown how burstiness curves can be used to characterize the burstiness of the traffic at different time scales. JPEG video exhibits no burstiness on small time scales. However JPEG video is bursty over longer time scales. MPEG video is bursty on short as well as long time scales. MPEG video is more bursty on small time scales than on large time scales. Traffic generated by NV is highly bursty over short periods of time but is very smooth when averaged over an interval of 10 seconds or more. Short term burstiness is a result of the coding algorithm and it's easier to handle. A traffic having short term burstiness can be made smoother at the cost of additional delays. Even if the traffic is not smoothed the worst case buffer requirements to serve such traffic at a constant bit rate is small. Long term burstiness is more difficult to handle. Most obvious way to handle long term burstiness is to design encoders which support constant bit rate operation.

An approach for wide area networks might be to reserve the bandwidth at peak bit rate for connections carrying video. Policing in this case would involve peak rate enforcement, but finding the peak rate for interactive video in advance may be difficult. The problem can be solved if the encoder limits the peak rate of its output stream.

Adjusting the sending rate depending upon feedback from the network can be a way to get better quality and congestion ratio. Network signals congestion to the application and based on this congestion signal the application adapts to a different rate by having a different compression ratio or fps ratio. This scheme may work well for interactive video but it is difficult for stored video because changing the bit rate of stored video may involve additional processing.