Bundle of Streams: Concept and Evaluation in Distributed Interactive Multimedia Environments

Pooja Agarwal, Raoul Rivas Toledano, Wanmin Wu, Klara Nahrstedt, Ahsan Arefin
Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, Illinois, United States
Email: {pagarwl, trivas, wwu23, klara, marefin2}@illinois.edu

***Abstract*—Distributed Interactive Multimedia Environments (DIMEs) enable geographically distributed people to interact with each other in a joint media-rich virtual environment for a wide range of activities, such as art performance, medical consultation, sport training, etc. The real-time collaboration is made possible by exchanging a set of multi-modal sensory streams over the network in real time. The characterization and evaluation of such multi-stream interactive environments is challenging because the traditional Quality of Service metrics (e.g., delay, jitter) are limited to a per stream basis. In this work, we present a novel cyber-physical, spatio-temporal QoS model that includes new metrics to measure and evaluate QoS for a bundle of streams and across bundles of streams. The new model allows better assessment of performance of any DIME application which employs multiple sensors. We validate our model by comparing the QoS performance of bundle of streams over three transport protocols - UDP, DCCP, and TCP in a 3D tele-immersive testbed.**

***Keywords*-Distributed Interactive Multimedia Environments, 3D tele-immersion, Quality of Service**

I. Introduction

Distributed Interactive Multimedia Environments (DIMEs) allow real-time collaborative physical activities among multiple, geographically distributed sites. DIMEs have been proven useful for many cyberphysical activities such as physical therapy, sport activities, and entertainment. Some of the specific real applications include Teleimmersive Dancing (TED) [12], remote training [8], virtual world gaming, and Tai Chi learning.

DIMEs are usually comprised of several sensing and actuating devices (e.g. cameras, microphones, body sensors, displays, haptic devices) as shown in Figure 1. At each site, sensing devices capture periodically current activities in DIME, forming a logical snapshot of the data produced by cameras, microphones, and body sensors. In general DIME model, each sensing device produces a different stream of data representing one dimension of the multimedia environment. These streams are then exchanged across several physical sites. Since all the streams capture pieces of the same scene, these streams show high spatial and temporal correlations. We call such cyber-physical spatio-temporal correlated streams a **Bundle of Streams**.

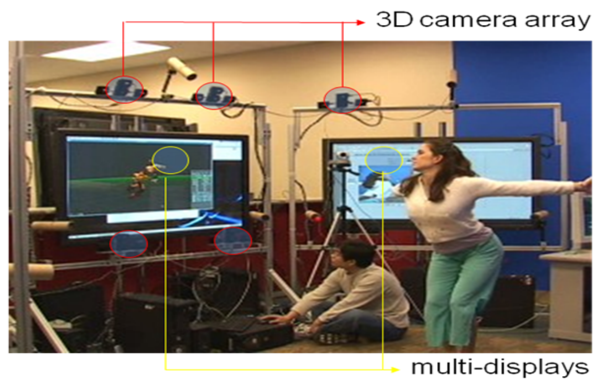


Figure 1. DIME Environment

In general to quantify the performance of distributed streaming, researchers use QoS (Quality of Service) metrics such as delay, jitter, loss rate. However, these traditional QoS metrics [6], [5], [4] work only on single independent streams(flows), thus, fail to quantify QoS for bundle of streams. In this paper, we propose a novel cyber-physical QoS model which includes new metrics to measure and evaluate QoS for a bundle and across bundles of streams. The bundle of streams occurs in various DIME environments such as teleimmersion and gaming. We also define two types of QoS metrics based on the correlations a) Cyber-physical QoS metrics for non prioritized bundle of streams and b) Cyber-physical QoS metrics for prioritized bundle of streams.

In evaluation, firstly we use our QoS model on a real DIME testbed (TEEVE) [14] and show how traditional metrics do not measure current DIME bundles correctly. Secondly, we use our metrics to compare the performance over three network protocols UDP, DCCP, and TCP in TEEVE and make a guided choice to select the best protocol for distributing bundle of streams across TEEVE sites.

In summary our contributions in this paper are:

1) Formalizing a concept of bundle of streams which exists in state-of-the-art collaborative applications;
2) A multi-dimensional cyber-physical QoS model which

is used to evaluate the bundle of streams in DIME environments;
3) Quatitative validation of better performance of our QoS metrics against traditional QoS metrics;
4) Evaluation of bundle performance over three transport protocols.

The rest of the paper is organized as follows: Section 2 gives an overview of the DIME model, Section 3 presents the theoretical QoS model for bundle of streams, Section 4 and 5 present the Cyber-physical non-weighted QoS metrics and weighted QoS metrics repectively, Section 6 presents the evaluation, Section 7 contains the related work and Section 8 is the conclusion.

II. Distributed Interactive Multimedia Environment Overview

Each physical space denoted as PS in DIMEs consists of a) users who perform tasks or roles b) sensors to capture the sensory information present in the environment c) display devices to present the captured sensory information to the users. Commonly, the sensors and display devices in a physical space are of different types, different ranges, and are placed at different positions inside the physical space. Typical sensors include cameras, microphones, human motion trackers, and wearable health monitors. Video display, speakers, and haptic devices are among the most typical display devices used in DIMEs. Thus, a physical space PS can be modeled as: Let the number of users, sensory devices, and display devices present in a physical space PS be denoted by $|U|$, $|SN|$, and $|DI|$ respectively, then we define a physical space as:

$$\begin{aligned} PS &= \{\{U_i, SN_j, DI_k\} | \forall\ i \in [0, |U|), \\ &\quad \forall\ j \in [0, |SN|), \forall\ k \in [0, |DI|)\} \end{aligned}$$

The sensors often produce data streams with heterogeneous frame sizes and rates. We denote the data stream produced by a sensor SN as S. The logical data unit of a stream is frame produced at time t (physical or logical), denoted as f_t^S. A stream is a sequence of frames in increasing order of timestamp as $S = (f_0^S, f_1^S, \ldots, f_n^S)$. Thus, a DIME environment is composed of several data streams being exchanged between the collaborating sites. All the notations used in the paper are also given in Table I.

To model the DIME environment, we map the physical spaces to a joint virtual cyberspace denoted by CS which consists of all the streams being produced by all the sensors in the physical spaces collaborating in a session SID. Thus, each sensor belonging to a physical space gets mapped to one stream in the cyber space. A cyberspace also maps the properties of each stream like frame rate, frame size, resolution which depend on the sampling rate and range of the sensor in the physical space. The correlations between the physical sensors are also mapped to the streams in the cyber space. Types of correlations observed between streams are as follows: (1) *spatial correlation:* The physical positions of sensors may have impact on the spatial correlation between their streams. For example, the video stream that captures the upper body of a user is correlated with the video stream that captures the lower body. (2) *temporal correlation:* The physical phenomenon being monitored also causes temporal correlation between consecutive observation of sensors and between sensory information captured by different sensors(synchronization issues).(3) *precedence correlation:* Some streams can be assigned higher priority depending on the type/importance of the streams. For example, the video streams which contain the side view of the user are less important than the video streams which contain the front view of the user.

To model a cyberspace CS, we map each sensor SN_j from each physical space PS_i collaborating at a given time t to a corresponding stream $S_j^{PS_i}$ in CS. Once all the streams are added to the CS, a correlation function $cor()$ is used to derive the correlation between all the streams belonging to CS. Let the number of physical spaces collaborating in a given session SID in CS be denoted by N_{SID}^{CS} and total number of sensors belonging to a physical space PS_i be $|SN^{PS_i}|$, then CS is defined as follows:

$$\begin{aligned} CS &= \{S_j^{PS_i} | \forall\ i \in [0, N_{SID}^{CS}) \\ &\quad \text{and } \forall\ j \in [0, |SN^{PS_i}|)\} \\ cor_{CS}(S_j, S_k) &= < spatial\_cor(S_j, S_k), \\ &\quad temporal\_cor(S_j, S_k), \\ &\quad precedence\_cor(S_j, S_k) >; \\ \forall\ j,k \in [0, |CS|) \end{aligned}$$

III. Theoretical Model

The QoS metrics defined in the literature like delay, jitter, loss rate are defined on a per stream basis and hence, cover only the relationships within a frame and between frames of a stream. They lack the ability to be applied to currently emerging cyber-physical streams which bear high spatial and temporal correlation with other streams as well as temporal correlation within themselves. Thus, for new collaborative appplications with several correlated integrated streams, we define a new QoS model.

A. Bundle of Streams

To define the QoS metrics on cyber-physical streams, we first define a mechanism to find groups of streams which need to be considered together given the spatial and temporal correlations. We introduce the concept of *Bundle of Streams* wherein, streams belonging to a common physical space and exhibiting high correlations belong to the same bundle. We present the definition of *Bundle of Streams* as follows:

Bundle of Streams*:A Bundle of Streams is a collection of highly correlated cyber-sensory streams captured in the*

Symbol	Definition
CS	cyberspace
PS	physical Space
S_i	stream i
S_i^{PS}	stream $i \in$ physical space PS
f_t^S	frame $\in$ stream S at time t
SN	a sensor
sid	session id
N_{SID}^{CS}	total number of physical spaces in CS in session SID
$\lvert SN^{PS_i}\rvert$	total number of sensors in the physical space PS_i
$\lvert CS\rvert$	total number of streams in the cyberspace CS
$cor_{CS}(S_j, S_k)$	function to calculate amount of correlation between two streams $\in CS$
$De(f_t^S)$	delay on frame f_t^S
$Ji(f_{ti}^S, f_{ti+1}^S)$	jitter between two frames $\in S$
$Sk(f_t^{S_0}, f_t^{S_1}, \ldots, f_t^{S_n})$	skew between frames across different streams
$LoI(S)$	total loss of information on S
$Numf_{LN}(S)$	total number of frames $\in S$ lost in network
$Numf_{LE}(S)$	total number of frames $\in S$ lost due to error
$Numf_{LD}(S)$	total number of frames $\in S$ lost due to expired deadline
$Numf_{LI}(S)$	total number of frames $\in S$ lost due to incomplete macroframe
B_{SID}^{PS}	bundle of streams originating in PS for DIME session SID
$threshold\_cor$	threshold on the amount of correlation required between two streams to belong to same bundle
$B_{SID}^{PS_1,\ldots,PS_n}$	bundle formed on bundle of streams originating in PS_i
$cor'_{CS}(B_{SID}^{PS_i}, B_{SID}^{PS_j})$	function to calculate amount of correlation between two bundles
$threshold\_cor'$	threshold on the amount of correlation required between two bundles to belong to a bundle of bundles
F_t^B	macroframe occuring at time t in bundle B
$\lvert BS\rvert$	total number of streams $\in$ bundle B
$\lvert F_t^B\rvert$	total number of frames $\in$ macroframe F_t^B
$De^B(F_t^B)$	bundle delay on macroframe F_t^B
$Ji^B(F_{ti}^B, F_{ti+1}^B)$	bundle jitter between two consecutive macroframes
$Sk^B(F_t^{B_0}, \ldots, F_t^{B_n})$	bundle skew between macroframes across different bundles
$LoI^B(B)$	bundle total loss of information on B
$Th^B(B, t_p)$	bundle throughput for bundle B with t_p as the unit time
t_p	time period for bundle throughput calculation
N_p	number of streams dealt as priority streams
$P(B, N_p)$	set of priority streams from bundle B
$WDe^B(F_t^P)$	weighted macroframe delay
$WJi^B(F_{ti}^P, F_{ti+1}^P)$	weighted macroframe jitter
$WSk^B(F_t^{B_0}, \ldots, F_t^{B_n})$	weighted skew between macroframes across sites

Table I
TABLE OF NOTATIONS.

same physical space PS during the same DIME session SID:

$$B_{SID}^{PS} = \{S_j^{PS} \mid cor_{CS}(S_i^{PS}, S_j^{PS}) \geq threshold\_cor \ \forall\ i, j \in [0, |SN^{PS}|)\}$$

The definitions of function $cor()$ and $threshold\_cor$ are application specific and depend upon the system parameters and types of sensors deployed in the system.

Example of Bundle of Streams: Figure 2 shows three 3DTI sites, PS_1, PS_2, and PS_3 collaborating with each other in session SID_t. As shown, PS_1 contains two cameras SN_1, SN_2, one audio device SN_3, and one body sensor SN_4. The bundle at site 1 is defined as $B_{SID_t}^{PS_1} = \{S_1^{PS_1}, S_2^{PS_1}, S_3^{PS_1}, S_4^{PS_1}\}$. It may be noted that in cases of low correlations, not all sensory streams produced at same physical space need to belong to the same bundle.

B. Bundle of Bundles

Apart from intra-site correlations between sensory streams, there exist further correlations between inter-site streams. Typical types of correlations between inter-site streams include temporal (synchronization and consistency issues) and precedence correlations (importance of one site over the other). Thus, we define a hierarchical bundle on top of the bundles received from different physical spaces as follows:

Bundle of Bundles: *A Bundle of Bundles is a collection of highly correlated bundles of streams belonging to different physical spaces $PS_1, \ldots, PS_n$ during same DIME session SID:*

$$B_{SID}^{PS_1,\ldots,PS_n} = \{B_{SID}^{PS_i} \mid cor'_{CS}(B_{SID}^{PS_i}, B_{SID}^{PS_j}) \geq threshold\_cor' \ \forall\ i, j \in [0, n)\}$$

In Figure 2, $B_{SID_t}^{PS_2,PS_3} = \{B_{SID_t}^{PS_2}, B_{SID_t}^{PS_3}\}$ shows an example of bundle of bundles defined across sites.

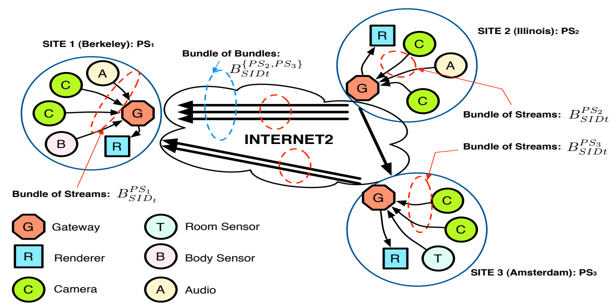


Figure 2. Representation of Bundle of Streams and Bundle of Bundles in a DIME architecture

C. Operations on Bundles

In fully functional DIME sessions, several operations need to be executed in physical and cyber spaces along various stages of a) initialization b) running c) maintenance and d) shutdown. These operations are necessary for the functioning of the DIME sessions, and define tasks to be performed on objects(e.g. session, stream, sensor, user) composing the DIME environment. For example common operations on a session include *AddUser*, *RemoveUser*, *QuerySession*, *AllocateResources*, and *MonitorResources*. As we are adding a new object *bundle* to the DIME environment, we cover set of operations that can be applied to a bundle. Using these operations a bundle can be initialized, modified, queried, and deleted. Common operations on a bundle are as follows:

1) $AddStream : B \times S \rightarrow B; AddStream(B, S) = B \bigcup \{S\}$ This operation allows adding new streams to a bundle.
2) $DeleteStream : B \times S \rightarrow B; DeleteStream(B, S) = B$ $\{S\}$ This operation allows deleting streams from a bundle.
3) $EditStream : B \times S \rightarrow B; EditStream(B, S) = edit(S, MD^S)$ This operation allows editing the stream and it's meta-data MD^S containing properties

and correlations like frame rate, size, resolution, spatial, temporal, and precedence correlations of a stream belonging to a bundle.

4) $QueryBundle : B \times Q \rightarrow S; QueryBundle(B, Q) = \{S_i : i \in |B|\}$ This operation allows querying a bundle with query Q and obtain set of streams which match the query.

D. Macroframes

After defining Bundle of Streams, we define a single cross-cutting entity on streams belonging to the same bundle on which the QoS metrics are applied. We define the concept of *Macroframe* as a logical snapshot of a bundle at time t. We formally define Macroframe as under:

***Macroframe**: A Macroframe is a set of frames/information representing a time-slice of a bundle at time t and/or within a small time interval $(t, t+\delta)$:*

$$F_t^B = \{f_{t_j}^{S_i} | \ \forall\ i \in [0, |B|); \ \forall\ t_j \in [t, t+\delta)\}$$

The number of frames composing a macroframe can differ in each macroframe depending on the relative rates of the streams and the time interval of the time-slice. For cases wherein a frame spans over more than one time-slice interval, it is considered to belong to only the macroframe defined on the earliest time-slice among them.

Example of Macroframe: In Figure 3, macroframe for time-slice t_1, $F_{t_1}^B = \{f_{t_1}^{S_1}, f_{t_1}^{S_2}, \ldots, f_{t_1}^{S_N}, f_{t_1+\delta}^{S_N}\}$

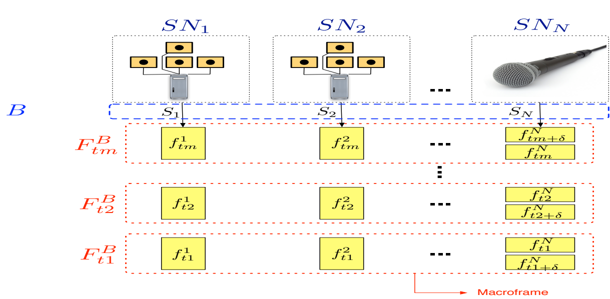


Figure 3. Representation of a Macroframe in a DIME architecture

IV. Cyber-Physical Non-Weighted QoS Metrics

Given the lack of QoS metrics for measuring the quality and performance in multi-dimensional correlated streams, we extend the current QoS model by defining cyber-physical spatio-temporal QoS metrics to handle integrated streaming. The cyber-physical spatio-temporal QoS metrics are unique in the sense that they are computed over two dimensions (both space and time) where spatial position of streams in cyber-physical space add spatial dimension and the temporal relationship between streams in cyber-physical space add the temporal dimension.

A. Bundle Delay

The DIME environments involve high level of user interactivity between different sites and hence, are very sensitive to overall delay experienced. In a bundle of streams, total delay not only comprises of the delay experienced on each stream in the bundle but also the time taken to complete a macroframe. Thus, delay for a bundle is a two-dimensional metric, with both temporal and spatial aspects in it.

***Bundle Delay:** For a given macroframe, the Bundle Delay De^B is the difference between the time at which the first frame belonging to the macroframe is sent at the sender and the time at which the last frame is received at the receiver as shown below:*

$$De^B(F_t^B) = max\ t_{recv}(f_i) - min\ t_{send}(f_i); \quad \text{where } 0 \le i < |F_t^B|.$$

Figure 4 shows an example of the bundle delay between two sites on bundle A with $De^B(F_i^A) = t'(f_n) - t(f_1)$.

B. Bundle Jitter

The DIME environments are composed of several sensory streams which are highly sensitive to jitter like audio and video streams. High jitter between macroframes can lead to poor overall quality. The formal definition of Bundle Jitter is as follows:

***Bundle Jitter:** For a given bundle, the Bundle Jitter Ji^B is the variation in $BundleDelay$ between two consecutive macroframes:*

$$Ji^B(F_{ti}^B, F_{ti+1}^B) = |De^B(F_{ti+1}^B) - De^B(F_{ti}^B)|$$

From Figure 4, for bundle A, the $Ji^B(F_i^A, F_{i+1}^A) = |De^A(F_{i+1}^A) - De^A(F_i^A)|$.

C. Bundle Skew

DIMEs consist of highly interactive real-time sessions among various users across geographically distributed areas, for example, playing virtual games to dance performances. In this, at any given instant t, a player/dancer needs to see a consistent view of all the remote sites, requiring synchronization of all the frames produced at time $t \pm \delta$ at all the sites. Thus, differences in delays across sites can reduce the responsiveness of the system. For such scenarios, the skew between macroframes across several sites is of importance.

***Bundle Skew:** Bundle Skew is defined as the maximum difference in $BundleDelay$ of macroframes belonging to different bundles of streams received from different sites:*

$$Sk^B(F_t^{B_0}, F_t^{B_1}, \ldots, F_t^{B_n}) = max|De^B(F_t^{B_i}) - De^B(F_t^{B_j})|; \quad \text{where } 0 \le i, j \le n, i \ne j.$$

From Figure 4, the $Sk^B = max(|De^B(F_i^A) - De^B(F_i^B)|, |De^B(F_i^A) - De^B(F_i^C)|, |De^B(F_i^B) - De^B(F_i^C)|)$

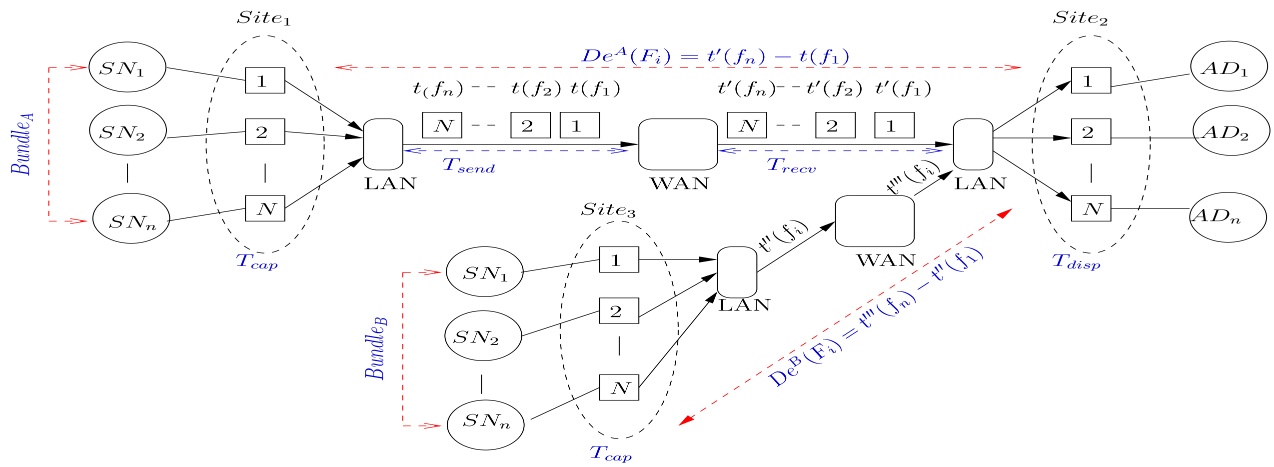


Figure 4. Bundle Metric Evaluation Framework

D. Bundle Loss of Information

Due to the large amount of data to be streamed across several sites, the DIME environments often face scarcity of network bandwidth. To adapt to the dynamics of the bandwidth fluctuations, DIMEs apply congestion control mechanisms. Common Application layer congestion control mechanisms for DIMEs include reducing frame rate by dropping frames, dropping non-priority streams, etc. Part of stream/frames can also get corrupted or lost while in transit. Both the above cases lead to loss of information from the receiver's perspective and hence, reduces the quality of the output at the receiver. Thus, we define bundle loss of information as follows:

Bundle Loss of Information: *The total Bundle Loss of Information comprises of (a) intermittent lost/dropped/delayed frames belonging to a stream (b) all the streams which are lost/dropped/delayed(in this, we consider all frames belonging to these streams as lost.) (c) frames which depend on the lost/dropped/delayed frames(belonging to same macroframe) in (a) and (b) and need to be dropped due to incompleteness. The unit of measurement is number of frames and bundle loss of information is defined as follows:*

$$LoI^B(B) = \sum_{0 \le i < |B|} Numf_{LN}(S_i) + Numf_{LE}(S_i) + Numf_{LD}(S_i) + Numf_{LI}(S_i)$$

E. Bundle Throughput

DIMEs are composed of highly dynamic heterogeneous streams with different frame rates across streams, variable number of bits per frame, and variable frame rate per stream. The granularity of the data is not at bit level but is at macroframe level i.e. end users see either the completed macroframe or none of it. Thus, the bundle throughput is defined as:

Bundle Throughput: *For a given bundle, the total size over the completed macroframes received per unit time t_p is the Bundle Throughput:*

$$Th^B(B, t_p) = \sum size(F_i^B)/t_p; \text{ where } t_{current} - t_p \le t_{recv}(F_i^B) < t_{current}$$

V. Cyber-Physical Weighted QoS Metrics

Observation 1 Based on the criticality of content, different streams can have different importance. For example, in [15], [13], the relative priorities are defined based on the contribution factor of streams towards the overall scene. The presence of prioritized streams directly influences how the QoS metrics need to be calculated. Any QoS disturbance on camera streams capturing the front view of the 3D model have greater impact on the overall quality of the rendered scene against QoS disturbance on camera streams capturing the side views.

Observation 2 Higher priority streams are allocated more resources to minimize delay, jitter, etc as compared to lower priority streams. Thus, priority streams may be rendered with faster rate. This leads to loose coupling between the frames.

Thus, depending on the relative importance of streams we define *weighted cyber-physical QoS metrics*. It is important to mention here that the inclusion of relative priorities of the streams add additional dimension to the spatio-temporal metrics defined in section IV, and hence, weighted cyber-physical metrics are truly a multi-dimensional metrics.

A. Weighted Bundle Delay

Based on the relative priorities given by $precedence\_cor(S_j, S_k)$, we bifurcate the bundle of streams into two sets a) High priority streams, and b) Low priority streams. For a given bundle, we pick first N_p number of streams from a list of sorted priority streams, where N_p is derived based on the characteristics of specific DIME and the dynamics of the system like current availability of resources. These N_p streams form a bundle of streams on which the bundle delay is computed as described below:

Weighted Bundle Delay: *For a given bundle, let* $P(B, N_p) = \{S_i^B | \ \forall i \leq N_p;$
$p(S_i^B) \geq p(S_{i+1}^B)\}$ *represent a bundle of priority streams; thus, the weighted bundle delay is the bundle delay experienced over the priority bundle* P*:*

$$WDe^B(F_t^P) \quad = \quad max \ t_{recv}(f_i) - min \ t_{send}(f_i); \quad \text{where } 0 \leq i < |F_t^P|.$$

B. Weighted Bundle Jitter

The weighted bundle jitter signifies the variation in the weighted bundle delay experienced in completing consecutive prioritized macro-frames:

Weighted Bundle Jitter: *For a given bundle, set of priority streams* P*, the weighted bundle jitter is the variation in weighted bundle delay between two consecutive macroframes belonging to* P*:*

$$WJi^B(F_{ti}^P, F_{ti+1}^P) \quad = \quad |WDe^B(F_{ti+1}^P) - WDe^B(F_{ti}^P)|$$

C. Weighted Bundle Skew

Similar to defining priorities for individual streams, priorities can be defined for individual sites based on the importance of the data being produced at the individual sites. These priorities can be defined on a system wide scale or on each receiver site.

Weighted Bundle Skew: *Weighted Bundle Skew is defined as the maximum difference in Weighted Bundle Delay of macroframes received from different sites:*

$$WSk^B(F_t^{B_0}, F_t^{B_1}, \ldots, F_t^{B_n}) \quad = \quad max|WDe^B(F_t^{B_i}) - WDe^B(F_t^{B_j})|; \quad \text{where } 0 \leq i, j \leq n, i \neq j.$$

VI. Evaluation

To evaluate our QoS model we use the TEEVE (TeleImmersion for Everybody) system, a real DIME system [14]. We first justify the need for new QoS metrics by evaluating the performance of traditional metrics in real DIME system. As a case study we use the metrics in our QoS model to compare performance of this TI system at the bundle concept under three transport protocols: Transmission Control Protocol (TCP), Datagram Congestion Control Protocol (DCCP)[7] and User Datagram Protocol (UDP).

A. Evaluation Setup

The TEEVE system is composed of multiple Teleimmersive (TI) sites. Each site is composed of cameras, displays and service gateways. Each stereo camera produces a 3D video stream with variable frame size in the range of 5KB to 30 KB, and variable frame rate ranging between 10 to 15 frames per sec. To make the experiments repeatable, we used a recorded creative dance performance. For each experiment we use 3 TI sites each at, University of Illinois at Urbana-Champaign (UIUC), Centrum Wiskunde & Informatica (CWI) in Amsterdam, and University of California at Berkeley (UCB). For the gateway server, we use 4 Dell Precision 670 with dual Intel Xeon processor at UIUC, DELL Optiplex 780 at UCB and AMD Athlon 64 Processor 3500+ at CWI.

B. Traditional QoS Metrics vs Cyber-physical QoS Metrics

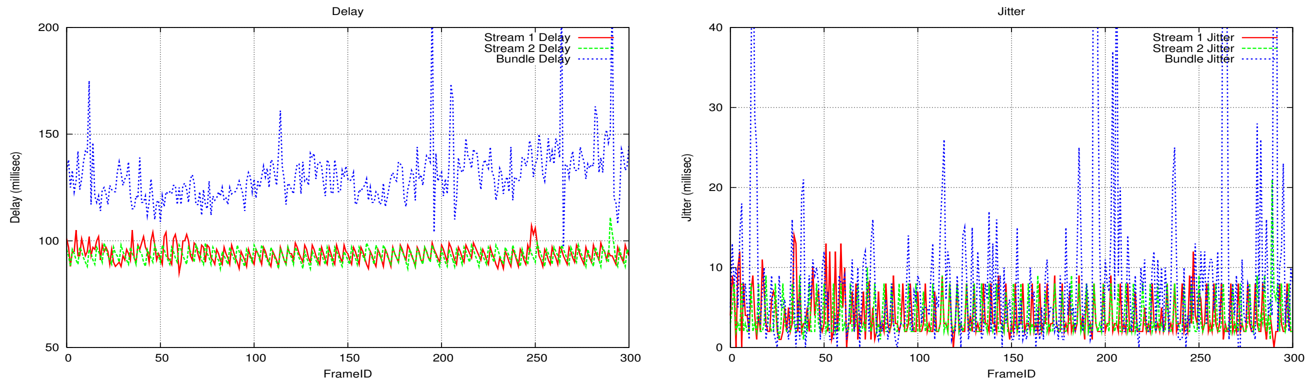
To justify the need of new QoS metrics, we evaluate the performance of traditional QoS metrics in DIME environment. A bundle consisting of two streams each is sent from CWI and UCB to UIUC. Figure 5(a) shows the delay experienced on each stream vs the delay experienced on a bundle. The delay on each stream only reflects a small portion(typically 55% to 65%) of the total delay experienced on a bundle. Thus, this signifies need of new metric to accurately quantify the delay DIMEs. Similarly, Figure 5(b) shows the inaccuracy of jitter calculated on each stream as compared to the actual jitter experienced on macroframes.

C. Case Study: Teleimmersive Systems

We apply our cyber-physical QoS metrics to evaluate the performance of the TEEVE system at UIUC. The goal is to evaluate the performance of TCP, DCCP, and UDP over bundle of streams framework. We choose these protocols as they are the most common choices for wide area real-time transmission. Next, we discuss the results as follows:

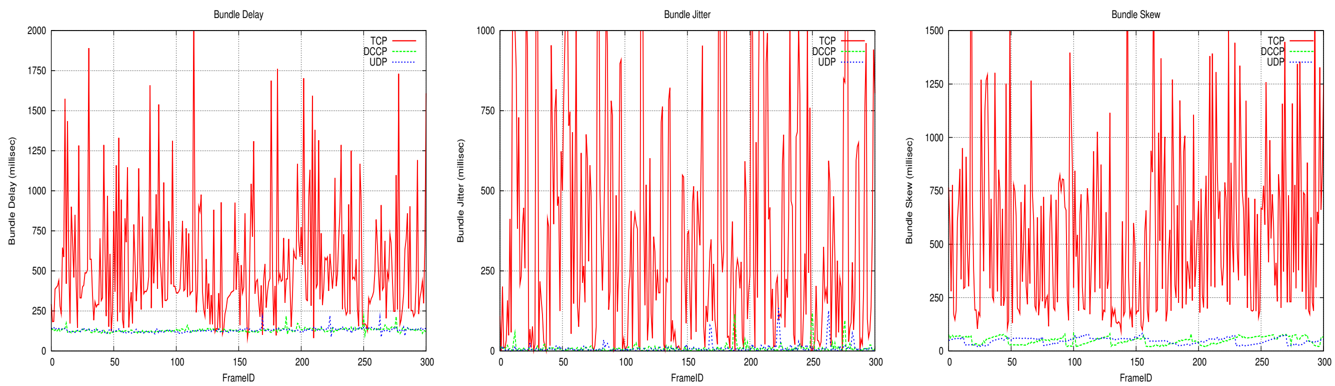
1) Experiment Set 1: In this set of experiments, non-weighted QoS metrics are used. Two bundles, one each from CWI and UCB are received at UIUC, and the number of streams in each bundle are varied between 2 and 3 streams. We performed the tests 6 times over different days and time, and the results show representative values of all of these tests. Due to space limitation, we show the results over CWI to UIUC connection, except for the bundle skew wherein skew between all three sites are shown. Figure 6 shows the QoS achieved over TCP, DCCP, and UDP in terms of bundle delay, bundle jitter, bundle skew, and bundle throughput. The bundle delay over TCP is significantly higher (about 500%) as compared to DCCP and UDP. Similar trend is seen on the bundle jitter on TCP, with values ranging from 10 milliseconds to 2000 milliseconds. The jitter on DCCP and UDP is very nominal ranging from 0 milliseconds to 100 milliseconds. This shows that the aggressive congestion control of TCP has high negative impact on the timeliness of the protocol.

The figure 6 shows the bundle skew on TCP (ranging between 150 to 1500 milliseconds) as compared to the skew on DCCP and UDP. For the evaluation of bundle loss of information, a stringent deadline of 200 milliseconds (5 fps) is defined on macroframe receipt(frame rates below 4 fps



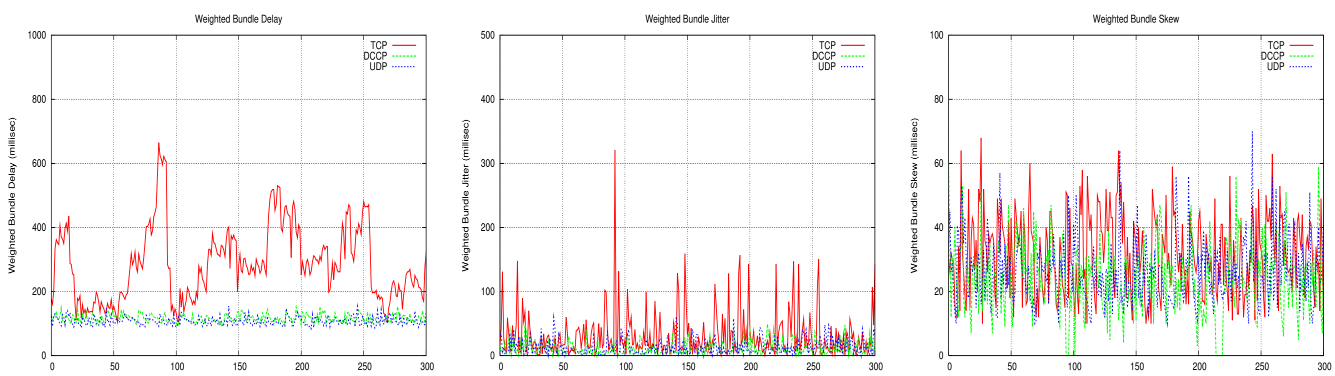
(a) Per Stream Delay vs Bundle Delay (over DCCP) (b) Per Stream Jitter vs Bundle Jitter (over DCCP)

Figure 5. Comparison of cyber-physical metrics against traditional metrics



(a) Bundle Delay (b) Bundle Jitter (c) Bundle Skew

Figure 6. Performance of TCP, DCCP, and UDP over cyber-physical metrics



(a) Weighted Bundle Delay (b) Weighted Bundle Jitter (c) Weighted Bundle Skew

Figure 7. Performance of TCP, DCCP, and UDP over weighted cyber-physical metrics

lead to complete unusability of DIME sessions [5]). The results on bundle loss of information show that 87% of macroframes are dealt as lost in TCP due to high delays, while only 1.8%, 3.2% are lost in DCCP and UCP(though they are unreliable transport protocols). With a source frame rate of 12 macroframes/sec, the bundle throughput attained over TCP ranges between 0 to 80 KBytes/sec, and between 20 to 180 Kbytes/sec over DCCP and UDP.

2) Experiment Set 2: This set comprises of evaluations done by using weighted metrics. Three camera streams constitute a bundle at each site, wherein 1 top camera and corresponding 1 bottom camera are given higher priority against third bottom camera. We also assign different rates to each camera stream depending on priority (15 fps, 10 fps, and 5 fps) to replicate the effects of variable macroframe sizes on the evaluation results. Figure 7 covers the obtained results. The weighted bundle delay on TCP shows significant reduction in delay as compared to the non-weighted bundle delay. Thus, the results indicate that one of DCCP/UDP can be used for wide area communication in DIMEs while TCP is prone to incur low QoS performance.

VII. Related Work

We cover the related work on QoS models and metrics present in the literature. [11] proposes consistency models for distributed multimedia applications with synchronization based approach. In [3], a QoS architecture for DIMEs is proposed based on levels of awareness across sites. In [2], authors propose a system wide QoS architecture spanning across all the layers of system and specify QoS mapping between the layers. The approach uses a synchronization mechanism called *'orchestration'* [1] between different streams to achieve synchrony. In [9], a framework for QoS mapping across different layers of a multimedia system is proposed. The framework spans over concurrent network, transport and application layer measurements of various QoS metrics.

Transport Modeling Research Group (TMRG) specifies per stream(flow) metrics in [6] . Wijesekera et al. in [5] [4], specify continuity and synchronization metrics. In [10], a multidimensional metric for evaluating perceptual quality of multimedia applications is proposed. To compare, our QoS model introduces a QoS framework for correlated set of streams called bundle of streams and we define multi-dimensional cyber-physical QoS metrics to evaluate distributed interactive systems.

VIII. Conclusion

This paper presents a new concept called "Bundle of Streams", along with a suite of QoS metrics to characterize and evaluate the emerging distributed interactive multimedia environments. The Bundle of Streams model captures the important spatio-temporal correlations among the multi-modal sensory streams in DIMEs, and serves as the conceptual foundation for a new set of QoS metrics such as Bundle Delay, Bundle Jitter, and Bundle Loss of Information. We provide formal definitions of these concepts and metrics, and validate them by comparing the performance of three transport layer protocols in a real DIME system.

IX. Acknowledgment

We thank Prof. Dick Bulterman, CWI and Ishan Vaishnavi, CWI for providing invaluable resources used in the experimentation. This research is supported by grants NSF CNS 08-34480, NSF CNS 07-20702, and Grainger Research Grant. The presented views are those of authors and do not represent the position of NSF.

References

[1] C. Andrew et al. A continuous media transport and orchestration service. *SIGCOMM Comput. Commun. Rev.*, 22(4):99–110, 1992.

[2] C. Andrew et al. Integrated quality of service for multimedia communications. In *Proc. IEEE INFOCOM'93*, pages 732–739, 1993.

[3] G. Chris et al. A qos architecture for collaborative virtual environments. In *Proc. ACM MULTIMEDIA '99 (Part 1)*, pages 121–130, New York, NY, USA, 1999. ACM.

[4] W. Duminda et al. Quality of service (qos) metrics for continuous media. *Multimedia Tools and Applications*, 3:127–166, 1996.

[5] W. Duminda et al. Experimental evaluation of loss perception in continuous media. *Multimedia Syst.*, 7(6):486–499, November 1999.

[6] S. Floyd. Metrics for the evaluation of congestion control mechanisms. Technical report, October 2007.

[7] S. Floyd et al. Profile for datagram congestion control protocol (dccp) congestion control id 3: Tcp-friendly rate control (tfrc). RFC 4342, 2006.

[8] K. Gregorij et al. Immersive 3d environment for remote collaboration and training of physical activities. In *Proc. IEEE VR*, pages 269–270. IEEE, March 2008.

[9] H. Jean-François et al. On end-to-end qos mapping. In *in Multimedia Networks, 21th IEEE Annual International Computer Software and Application Conference (COMPSAC)*, pages 303–314, 1997.

[10] C. Mark et al. End-to-end quality in multimedia applications. In *Chapter 40 in Handbook on Multimedia Computing, CRC Press, Boca*, 1998.

[11] X. Qin. A delayed consistency model for distributed interactive systems with real-time continuous media. *Journal of Software*, 13(6):1029–1039, June 2002.

[12] S. Renata et al. Advancing interactive collaborative mediums through tele-immersive dance (ted): a symbiotic creativity and design environment for art and computer science. In *Proc. ACM Multimedia '08*, pages 579–588. ACM, 2008.

[13] B. Yu. *MyView: Customizable Automatic Visual Space Management for Multi-Stream Environment*. PhD thesis, University of Illinois at Urbana-Champaign, September 2006.

[14] Y. Zhenyu et al. Teeve: The next generation architecture for tele-immersive environment. In *Proc. IEEE ISM '05*, pages 112–119. IEEE Computer Society, 2005.

[15] Y. Zhenyu et al. Enabling multi-party 3d tele-immersive environments with viewcast. *ACM TOMCCAP*, 6(2):1–30, 2010.