Dante: An FOV-Aware FEC-Based Multipath Protocol for 360-Degree Video Streaming

Abstract

360-degree videos have been widely applied due to its unprecedented immersive experience. The appreciation of watching 360-degree videos on untethered mobile devices, such as smartphone headsets, are considered to be a more promising trend, which, however, is hampered by the poor Quality of Experience (QoE), subject to limited bandwidth and time-varying characteristic of wireless networks. Fortunately, the key observation that only a small portion of a 360-degree video is perceived by users at any time, i.e., Field Of View (FOV), implies that unequal attention should be given different regions spatially and can be utilized to mitigate dependence on stable network condition and ultra-high bandwidth. Currently, the state-of-the-art schemes, like FOV-aware tile-based streaming in the application layer, almost based on FOV-aware bit-rate adaptation, transport protocol of which, however, fails to consider FOV. So, we propose an application-layer protocol, reliability scheme of which is FOV-aware and performs hierarchical protection to boost QoE of 360-degree videos in mobile scenarios. Experiments demonstrate our protocol achieves desirable improvements over the reference schemes in QoE of 360-degree videos.

1. INTRODUCTION

With the promise of immersive visual experience, 360-degree videos are widely applied in sports field, social field [1], and apps development field [2], etc. Meanwhile, it has been well deployed across mainstream content providers such as NBC (who broadcast the 2018 winter Olympics in VR), news outlets, such as CNN, New York Times, and user-generated content platform such as Youtube and Facebooks. However, some significant challenges still remain as the major bottlenecks for 360-degree technologies, even worse in mobile scenarios.

The bottleneck can be summarized as that wireless links are characterized by limited resources and time-varying conditions, while 360-degree video streaming is characterized by bandwidth intensiveness and delay sensitivity.

For example, application round-trip latency is supposed to be less than 10ms for imperceptible Motion-To-Photon(MTP) latency in order to prevent simulator sickness [18], due to the refresh rate of Head-mounted Display(HMD), such as 120 Hz. Such the rigorous delay constraint prevents the implementation of those traditional protocols based on feedback and retransmissions.

Meanwhile, 360-degree videos have high bandwidth requirement in order to enable immersive experience. The bit-rate of 8K 360-degree videos at 60 fps encoded using High Efficiency Video Coding (HEVC) [22] is around 100 Mbps. However, according to Opensignal [3], covering 77 countries in the world, almost a half of countries have access to 4G cellular network with only 10-25 Mbps speed in 2017. Obviously, it's not available and practical to deploy 360-degree videos service in mobile scenarios.

Currently, FOV-aware tile-based streaming, a scheme of application layer, is proposed to mitigates bandwidth and delay, due to the introduction bit-rate adaptation and prefeatching video segment. However, it can not perform well over wireless links because the most advanced transport protocols it uses are not FOV-aware, for example, their reliability schemes of transport layer fail to consider FOV.

In this paper, we propose Dante, as depicted in Figure 1, an application-layer 360-degree video protocol, the reliability scheme of which is FOV-aware and can perform hierarchical protection spatially to boost QoE of 360-degree videos in mobile scenarios.

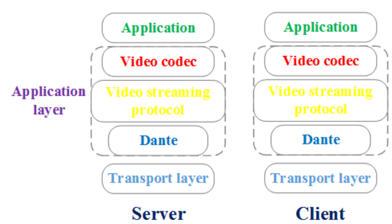


Figure 1: Illustration Of Dante

2. BACKGROUND AND MOTIVATION

2.1 360-degree Tile-based Streaming

In order to mitigate dependence on high bandwidth, FOV-aware tile-based streaming scheme, based on Dynamic Adaptive Streaming over HTTP (DASH) [14], is extensively studied in recent years. For most typical works [8][25][19] [10][12][16],

Solution	protocol layer	video distortion model	data recovery	adaptive FEC	FEC parameter decision	data allocation	video-awareness	FOV-awareness
MPLOT[21]	transport layer	×	RS codes & retransmissions	✓	balance between goodput and recovery probability	packet generation's order	×	×
FMTCP[9]	transport layer	×	Raptor codes & retransmissions	✓	balance between goodput and recovery probability	packet delivery time minimization	×	×
MPMTP[15]	application layer	×	Raptor codes	✓	goodput maximization	block arrival time minimization	×	×
CMT-VR[24]	transport layer	inflexible model	Raptor codes & retransmissions	✓	utility maximization	utility maximization	✓	×
Dante	application layer	proposed flexible model	FFT-RS code	✓	hierarchical protection and distortion minimization	hierarchical protection	✓	✓

Table 1: MAIN DIFFERENCE OF THIS WORK WITH THE EXISTING WORKS

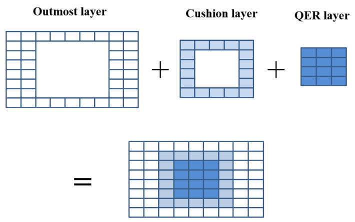


Figure 2: Video Representation

360-degree videos are split into segments of equal length, such as 1s, and the server offers multiple bit-rate of representations for every segment. Furthermore, with the technology of motion-constraint tile sets (MCTS), every segment can be encoded into multiple tiles spatially, as depicted in Figure 1, each of which can be independently decoded, stored into a single file and sent to clients alone. So, the server offers representations that also differ spatially by having a Quality Emphasized Region (QER) [8]: a region of the video which is made up of tiles with higher bit-rate than the rest of tile of the remaining of the video. Clients periodically pre-fetch a representation for the next segment such that the bit-rate adapts to available bandwidth and QER best matches the expected viewport of users. Unfortunately, they only strive to optimize FOV-aware bit-rate adaptation schemes in application layer and failed to consider a effective FOV-aware scheme to guide the transport protocol. Thus, they can not perform well over lossy links, even worse in mobile scenarios.

2.2 MultiPath Parallel Transmission With FEC

MultiPath Parallel Transmission, considering mobile devices, like smartphone, almost equipped with diffrent radio interfaces (eg.Wi-Fi and LTE), is considered to be a promising way to solve the problem of limited bandwidth over wireless links. IETF-MPTCP [6] is proposed and suffers from the performance degradation subject to the bottleneck link. These works [21][9] [13], due to the introduction of FEC and well designed data allocation algorithm, not only aggregate capacities across paths but counter wireless network's time-varying characteristic, mitigating the head-of-line blocking and packet out-of-order in multiple diverse network.

the intorduction of FEC, mitigating retransmissions, is a important scheme to reduce delay and strengthen the scheme of realibilty guarateen. FEC[1] is introduce to mitigate retransmissions, greatly achieving low delay and high throughput. Considering a chunk of data organized into K packets, with equal length, the FEC encoder takes the K data packets, and adds redundant M FEC packets to create a coded block of size $\mathrm{B} = (\mathrm{K} + \mathrm{M})$. The eceiver can recover completely the origin data of K packets if any at least K packets of $K + M$ packets are received. The code rate is equal to $K/(K + M)$. Obviously, it can counter M packets of loss over lossy links at most and when code rate is smaller, i.e., more redundant FEC packets, it has more powerful recoverability.

Obvious, the data of FOV region, which is be viewed by users with a higher probability, is more imporatant than non-FOV data. However, even combined with 360-degree tile-based streaming, due to the failure to consider FOV, the state-of-the-art transport schemes, which are not FOV-aware, spend as same amount of bandwidth on reliable delivery of trival data, i.e., non-FOV data, as FOV data. So, can we, from the perspective of reliablity, design a protocol, which prioritizes the reliability of FOV data over non-FOV data, by sacrificing the quality of non-FOV data.

So, unlike the 360-degree tile-based streaming protocol, which makes the tiles expected viewed by the users streamed with higher bit-rate and the rest of tiles with low bit-rate to boost QoE of 360-degree video, we proposed Dante, realiability scheme of which performs on different region of videos in a hierarchical fashion, i.e, preferentially provisioning the data closer to the FOV region with more redundancy of FEC. Thus, the data, which more strongly affect the QoE of video, transimitted over lossy links, is supposed to integrally received with a higher probability and thus QoE of 360-degree video can be boosted notably.

Manwhile, in Table 1, we summarize the main differences of Dante with the existing multipath schemes. To the best of our knowledge, Dante is the first FOV-aware 360-degree video protocol over heterogeneous wireless networks.

3. PROTOCOL DESIGN

3.1 System Overview

[1]In Dante, FEC is generated through erasure coding of blocks of packets, to recover lost packets, which should be distinguished from FEC, computed on individual packets using channel coding, recovers from bit errors.

Dante is proposed to support high-quality 360-degree video streaming service over the wireless network. In Dante, UDP combined with Systematic FEC, FFT-RS code [17], which enjoys ultra-high throughput and low coding delay, is integrated to provide reliable delivery over wireless networks. Besides TCP is supplementary to exchange control packets, which are of significance. The proposed Dante is implemented at a shim layer between video adaptive streaming scheme and transport layers. According to [20], 360-degree videos have less variability than traditional videos, researchers can pay more attention to the new requirement of the network responsive to user changing the field of view. So in this paper, we assume that video contents are encoded with a constant bitrate, The overall protocol architecture is illustrated in Figure 2.

At the server side, the key components include a network status monitor, a parameters controller and some executives. The network status monitor receives channel's status feedback from the client, including RTT, packet loss rate, available bandwidth, and delivers them to the parameter controller, which is composed of FEC parameters adjustor (Algorithm 1) and packet scheduler(Algorithm 2), and makes decisions about the FEC parameter and the packet scheduling vector, periodically. The video data, first encoded by FFT-RS encoder, is then sent over the path according to the packet scheduling result.

At the client side, Dante reorders packets received before the deadline, and examine the number of original data symbols in the current block. If any data symbol lost in packets, the FFT-RS decoding is applied to recovery lost packets. Otherwise, video data packets are directly delivered to the video codec, avoiding unnecessary decoding. Due to systematic feature of FFT-RS, despite decoding failed because of no enough redundant data packets, the system still can obtain as many source symbols as possible, then passing them on to the upper application.

Congestion control, the combination of TCP-friendly rate control (TRFC) [11] and Spike [7] is adopted in order to regulate the transmission rate of UDP sub-flows, which guarantees the fairness with other sessions and mitigates unnecessary performance degradation in throughput. **video bitrate adaption,** our protocol just takes a scheme, like DASH [14], selecting preferentially higher bitrate for the region of video closer to the expected viewport, to adapt to the available bandwidth.

3.2 Model Description

The aim of protocol is to find the optimal set of action in order to maximize the QoE. Video distortion can be considered as a standard metric of QoE. So, if we can design a 360-degree videodistortion computing model, which can not only manifest the expected effect of tile-based bitrate selection on QoE, but the effect of FEC provisioning on QoE. Then, utilizing the model to instruct every action of protocol, we can greatly boost the QoE of 360-degree video.

3.2.1 Video Representation And Video Distortion Model

Video distortion model. However, Traditional video distortion models, like computing SSIM(Structural Similarity), rigidly attach equal importance to every pixel of frames. Generally, they can be formulated as [24]: $d_m = d_{m,trunc} + d_{m,drift}$ where $d_{m,trunc}$ denotes truncation distortion depicting the expected level of packet loss after FEC provisioning and $d_{m,drift}$ denotes the drift distortion, depicting the expected level of losses of parent frames, due to video decoding dependencies in a GOP. Detail derivation is seen in paper[24]. That is not appropriate to be applied in the 360-degree video, where only a small region for videos is effective and perceived by users at any time. That implies a flexible video distortion model is supposed to be considered.

And we propose a distortion model based on the combination of two distortion models from the paper [25] and the paper [24]. As mentioned earlier, only a small portion of the 360-degree video is watched by users at any time, which implies unequal attention needs to be given different regions spatially. So, we attach a weight value to the distortion for every layer. And thus, total distortion for every frame is divided into three parts, for each of which, the distortion can be formulated as following:

$$d_m^\alpha = \gamma^\alpha (d^\alpha_{m,trunc} + d^\alpha_{m,drift})$$

Therefore the distortion for every frames in 360-degree video is formulated as:

$$d_{m,effetive} = \sum_{\alpha \in Q} \gamma^\alpha (d^\alpha_{m,trunc} + d^\alpha_{m,drift})$$

where $d_{m,effective}$ denotes effective distortion of the m-th frame. According to 360ProbDASH[25], each tile of 360-degree videos is expected to be watched by users with a probability following Gaussian Distribution at any time, which implies γ^α can be abtained with the probabilistic model of 360ProbDASH:

$$\gamma^\alpha = \sum_{i=1}^{\Omega^\alpha} p_i \cdot S_i \tag{1}$$

Where p_i stands for viewing probability of the i-th tile in the α layer, S_i denotes spherical area of the i-th tile and Ω^α denotes tiles set of the corresponding layer. The detail derivation is seen in paper [25]. So, for every frame, the distortion is formulated as following:

$$d_{m,effective} = \sum_{\alpha \in Q} \sum_{i=1}^{\Omega^\alpha} (p_i \cdot S_i (d^\alpha_{m,trunc} + d^\alpha_{m,drift}))$$

where $d^\alpha_{m,trunc}$ and $d^\alpha_{m,drift}$ are derivated the same way as performed in the traditional distortion model aforementioned [24].

3.3 Algorithm Design

Compared to traditional FOV-agnostic protocols, our protocol can further effectively utilize available network resource

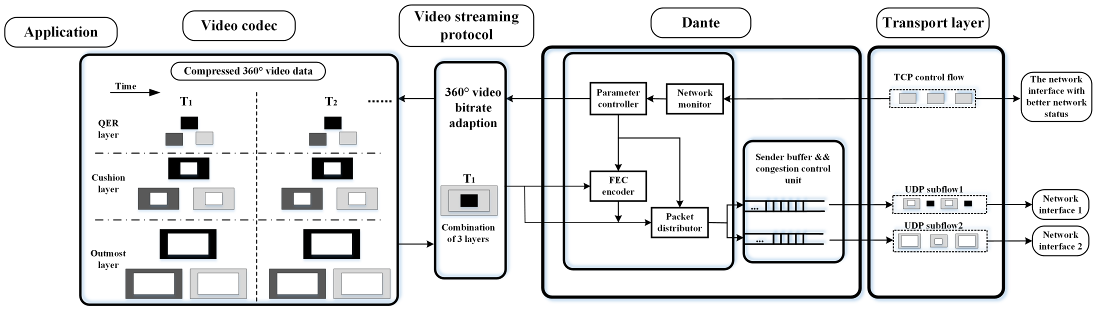


Figure 3: The Architecture of Protocol

due to FOV-aware benefit. To achieve this, 360-degree video features including layer priority, frame priority are integrated in our protocol. We consider the expected video quality loss caused by transmission loss and decoding dependencies of video codec, both of which depend on the FEC parameter adjusting procedure.

3.3.1 FEC Code Rate Adaptive Adjusting

Unlike the video bit-rate adaptation, Dante, from the perspective of reliability scheme, preferentially provisions the tiles, which are viewed by user with higher probablity and are more important to user's QoE, with more FEC redundancy.

However, increasing the FEC redundancy can mitigate the transmission loss, but, with the introduction of computing overhead [23], over-provisioning of FEC may enlarge the end-to-end delay, even causing unnecessary data loss caused by the hit of deadline. Assuming after packet scheduling, seen as in section 3.4.2, our system has attained the packet scheduling action vector, $\Phi_m^\alpha\{\alpha \in Q, 1 \leq m \leq M\}$ by Algorithm 2, imposed on the m-th frame in α layer, which can be formulated as following: $\Phi_m^\alpha = p, \quad \text{if scheduled} \quad \text{to path p}$. So, the code rate should be carefully adjusted. We consider the minimization of distortion as the goal of our protocol's FEC parameters adjusting, subject to constraints of deadline and available bandwidth. So, the optimization problem can be formulated as following:

$$\{R_m^a\}_{1 \leq m \leq M, a \in Q} = \arg\min(\sum_{i=1}^{M} d_{effective}^m). \tag{2}$$

$$\text{subject to} \quad T_m^{enc} + T_m^{dec} + T_m^{tran} \leq T, \quad 1 \leq m \leq M, \tag{3}$$

$$\text{and} \qquad \lambda^p(\Phi) \leq \mu_p \quad , \qquad 1 \leq p \leq P, \tag{4}$$

$$T_m^{enc} = \eta_{enc} \cdot s \cdot \sum_{\alpha \in Q} n_{m,\alpha} \cdot \log k_{m,\alpha}, \tag{5}$$

$$T_m^{dec} = \eta_{dec} \cdot s \cdot \sum_{\alpha \in Q} n_{m,\alpha} \cdot \log n_{m,\alpha}, \tag{6}$$

$$T_m^{tran} = \frac{V_m}{\frac{E_S\sqrt{1.5}}{R\sqrt{p}}} \ , \tag{7}$$

$$\lambda^p(\Phi) = \lambda \frac{\sum_{m=1}^{M} \sum_{\alpha \in Q} V^\alpha_{m\{\Phi_m^\alpha \in p\}}}{\sum_{m=1}^{M} \sum_{\alpha \in Q} V_m^\alpha}. \tag{8}$$

where, $\{R_m^\alpha\}$, denotes the code rate in the m-th frame for the α layer. The first constraint indicates that, due to video's timeliness, video data is supposed to be received before the playback deadline, and delay constraint T is imposed on every frame of GOP to restrict the downloading time of each frame, the sum of FEC encoding time, decoding time and transmission time, supposed to be smaller than T. Consequently, it's impractical to derive the optimal solution with polynomial time complexity, and the greedy search is not applicable for real time applications. To solve this problem, we design a fast research algorithm, which complexity is $O(N \cdot M \cdot Q)$, to obtain a sub-optimal solution of FEC code rate adaptive problem, shown in Algorithm 1.

Algorithm 1 FEC code rate adaptative algorithm

```
R = min(Eq. (3), Eq. (4)) , according to delay constraints, Eq. (3) and
    bandwidth constraints Eq. (4),
for α ∈ Q do
    Calculate γ^α , according to Eq. (1),
end for
N = (V/S) · R
for i = 1toN do
    index = 0, Δ_d = 0
    for m = 1toM do
        for α ∈ Q do
            d_effetive = Σ_{0≤m≤M} Σ_{α∈Q} γ^α (d^α_{m,trunc} + d^α_{m,drift})
            N_{m,α} = N_{m,α} + 1
            Δ = | −d_effective + Σ_{0≤m≤M} Σ_{α∈Q} γ^α (d^α_{m,trunc} + d^α_{m,drift}) |
            N_{m,α} = N_{m,α} − 1
            if Δ ≥ Δ_d then
                index = m,    Δ_d = Δ
            end if
        end for
    end for
    N_{m,α} = N_{m,α} + 1
end for
return { R_{m,α} = K_{m,α} / N_{m,α} }_(1≤m≤M, α∈Q)
```

3.3.2 Packets Scheduling

Our protocol allocates preferentially the better bandwidth resource to the region closer to FOV. Besides, the video data from the region closer to FOV layer will be put into a path with the higher rank value under the circumstance of the constraint of delay and available bandwidth. And the rank value of path is calculated considering the transmission rate presented in the paper[11]. The algorithm is described in Algorithm 2.

Algorithm 2 Packet scheduling algorithm

for α in Q (QER layer, cushion layer, outermost layer) **do**
  for m = 1 to M **do**
    calculate ranks, $Rank^x = \frac{n_{buf_x}^{total} \cdot S}{tr_x^{TFRC}} + \frac{K_m^{\alpha} \cdot S}{tr_x^{TFRC}(1-p^x)} + t_{ott}^x$;
    find the path x_{tmp} with lowest rank value;
    $n_{buf_x}^{total} \Leftarrow n_{buf_x}^{total} + K_m^{\alpha}$;
    if $Rank_{x_{tmp}} \leq T_S$ And $\frac{n_{buf_x}^{total}}{tr_x^{TFRC}} \leq u_x$ **then**
      $\Phi_m^{\alpha} = x_{tmp}$;
    else
      $n_{buf_x}^{total} \Leftarrow n_{buf_x}^{total} - K_m^{\alpha}$;
    end if
  end for
end for
return $\Phi_m^{\alpha}\{1 \leq m \leq M, \alpha \in Q\}$;

Packet scheduling action Φ_m^{α} imposed on the m-th frame in α layer can be formulated as following:

$$\Phi_m^{\alpha} = p, \quad \text{if scheduled to path p} \ .$$

4. PERFORMANCE EVALUATION

We conducted extensive experiments to demonstrate Dante performance gains on QoE. Specifically, the evaluation metrics considered in this work include single scale and multi-scale structural similarity(SS-SSIM and MS-SSIM), Signal-to-Noise Ratio(VSNR), PSNR and frame loss frequency.

4.1 Experimental Set-up

Experiment topology: 5 sources connect to node N0, while 5 sinks connect to node N1. And N0 connects to N1 through two lossy links. Sources and sinks represent servers and clients respectively, and Dante is deployed on all servers and clients. The topology is not shown because of space limitation. The experiment scenario is that sources send the data to sinks through two lossy links with the request of video segmentations.

Testbed configuration: The sources and sinks are commodity servers with Ubuntu 16.4 (kernel 4.40), each of which is equipped with an Intel(R) Core(TM) i3-4150k cpu @ 3.5GHz (4 cores), two Intel 82599ES 10G dual port NICs and 32 G memory.

360-degree video content preparations: We download three 8K video sequences from Youtube, each of which is transcoded into three kinds of bitrate, such as 50Mbps, 12Mbps and 3Mbps, and then cut into 60 video segmentations with 1s length. In this way, we can get 180 files for every video sequence. At the last step, we utilize 360transformation [5] to crop spatially each of these files into three layers including the QER layer, cushion layer and outmost layer, and total 540 video representations for every video are obtained. All files of video segmentations are preprocessed, generated and managed on the server side.

Network parameter set: Gilbert model is adopted to mimic the packet loss pattern in real wireless networks, supported by traffic control (TC) [4], in which four parameters(ξ_i^G, ξ_i^B, 1-h and 1-k) are needed, ξ_i^G and ξ_i^B are transition probabilities between the bad and good state, 1-h and 1-k is the loss probability in the bad state and good state, respectively. Meanwhile, average packet loss rate is equal to $\pi_i^B = \xi_i^B / (\xi_i^B + \xi_i^P)$. Meanwhile, the bandwidth of two path is set randomly in corresponding specified range and RTT is set as 50ms and 100ms for Path-A and Path-B, respectively. Besides, 1-h and 1-k are set as 1 and 0, respectively. Parameter set is described in detail in Table 2.

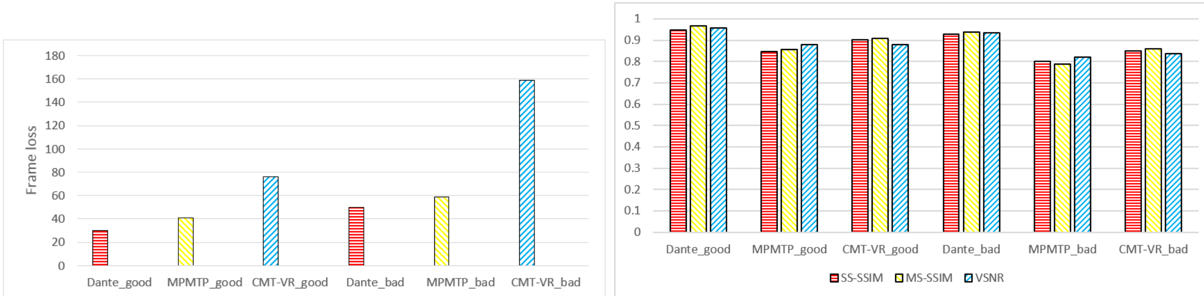
(A) Network	Time(Sec.)	Bandwidth(Mbps)	RTT(ms)	Pakcet loss rate(%)
Path-A	0∼20	15∼40	50	ξ_i^B = 0.03, ξ_i^G = 10
Path-A	20∼40	15∼40	50	ξ_i^B = 0.05, ξ_i^G = 10
Path-A	40∼60	15∼40	50	ξ_i^B = 0.1, ξ_i^G = 10
Path-B	0∼20	10∼20	100	ξ_i^B = 0.05, ξ_i^G = 10
Path-B	20∼40	10∼20	100	ξ_i^B = 0.1, ξ_i^G = 10
Path-B	40∼60	10∼20	100	ξ_i^B = 0.3, ξ_i^G = 10
(B) Network	Time(Sec.)	Bandwidth(Mbps)	RTT(ms)	Pakcet loss rate(%)
Path-A	0∼20	15∼40	50	ξ_i^B = 0.05, ξ_i^G = 10
Path-A	20∼40	15∼40	50	ξ_i^B = 0.1, ξ_i^G = 10
Path-A	40∼60	15∼40	50	ξ_i^B = 0.2, ξ_i^G = 10
Path-B	0∼20	10∼20	100	ξ_i^B = 0.15, ξ_i^G = 10
Path-B	20∼40	10∼20	100	ξ_i^B = 0.3, ξ_i^G = 10
Path-B	40∼60	10∼20	100	ξ_i^B = 0.5, ξ_i^G = 10

Table 2: Network Condition of Two Wireless networks: (A)Relatively Good Wireless Conditions And (B)Relatively Bad Wireless Conditions

4.2 Performance Comparison With Existing Protocols

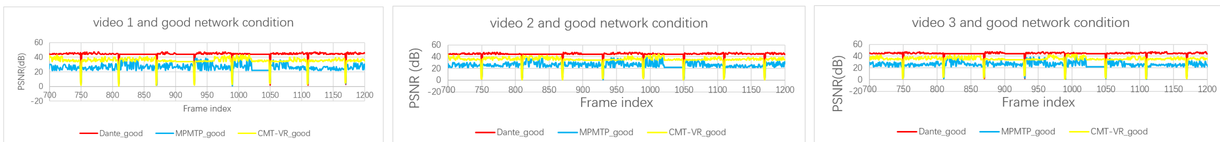
We consider MPMTP [15], content-agnostic, and CMT-VR [24], content-aware, as reference schemes, both of which utilize Raptor codes to guarantee reliability, mitigating the unnecessary retransmissions.

First, as show in Figure 3.a, Dante encounters the minimum times of frame loss event no matter in relatively good or bad network condition. Meanwhile, to demonstrate the perceived 360 degree video quality, Figure 3.b gives a group of results from multiple metrics including SS-SSIM, MS-SSIM and VSNR, we can see that Dante, compared with MPMTP and CMT-VR, better performs in subjective video quality in good network status or relatively bad network status. Then, Figure 4 and Figure 5 compare instantaneous PSNR of three video sequences in good network condition and bad network condition, respectively. The result shows that Dante achieves a great improvement of PSNR, compared to MPMTP and CMT-VR. The reason why MPMTP performs worst in all protocols is that, despite no involving retransmissions and maximizing the throughput, it doesn't consider video's inherent feature, such as decoding dependencies of video codec, which should have give different frame unequal attention, and thus can't utilize effectively the network allocation to boost video quality. Meanwhile, CMT-VR performs better than MPMTP, due to its consideration of frame priority. However, Unfortunately, not involving any feature of 360 video feature makes it waste valuable bandwidth on unimportant data, thus CMT-VR is the secondary one. Dante takes into acount not only traditional video fea-



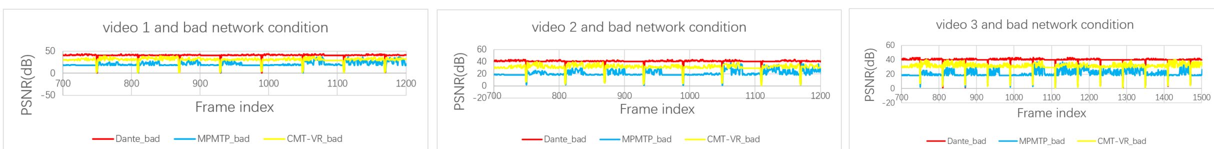
(a) Average Number Of Frame Loss Event (b) Multiple Metrics Of Video Quality

Figure 4: Average Frame Loss Number And Multiple Metrics For video Quality In Good And Bad Network Condition



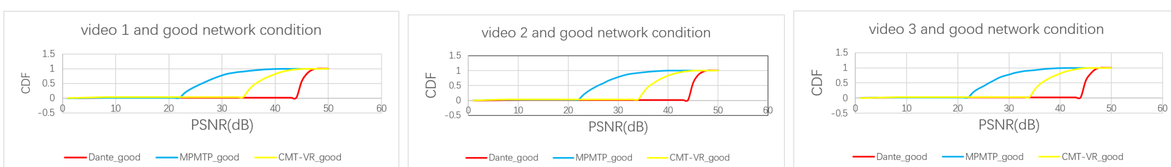
(a) Video Sequence 1 (b) Video Sequence 2 (c) Video Sequence 3

Figure 5: Instantaneous PSNR In Relatively Good Network Condition



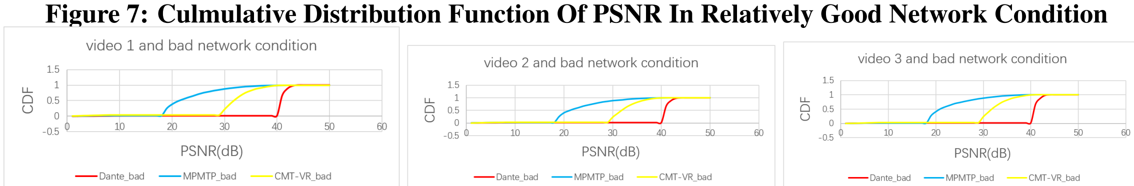
(a) Video Sequence 1 (b) Video Sequence 2 (c) Video Sequence 3

Figure 6: Instantaneous PSNR In Relatively Bad Network Condition



(a) Video Sequence 1 (b) Video Sequence 2 (c) Video Sequence 3

Figure 7: Culmulative Distribution Function Of PSNR In Relatively Good Network Condition



(a) Video Sequence 1 (b) Video Sequence 2 (c) Video Sequence 3

Figure 8: Culmulative Distribution Function Of PSNR In Relatively Bad Network Condition

tures aforementioned, but the 360-degree video characteristic, such as FOV. Benefiting from the hierarchical protection spatially and temporally, Dante achieves desirable upgradation in instantaneous PSNR and cultimulative distribution function of PSNR shown in Figure 6 and Figure 7.

5. CONCLUSION

In this paper, we propose a multipath protocol for 360-degree videos, which, based on multipath and FEC, performs hierarchical protection to counter limted bandwidth and time-varying problems in wireless networks. Experiments demonstrate Dante achieves desirable gains on 360-degree video QoE against reference schemes.

ACKNOWLEDGEMENT

The work was supported by the National Key Basic Research Program of China (973 program) under Grant 2014CB347800, National Key Research and Development Program of China under Grant 2016YFB1000200 and the National Natural Science Foundation of China under Grant No. 61522205, No. 61772305, No. 61432002. Dan Li is the corresponding author of this paper.

6. REFERENCES

[1] https://facebook360.fb.com/.

[2] https://developers.google.com/vr/discover/360-degree-media.

[3] https://opensignal.com/reports-data/global/data-2017-11/report.pdf.

[4] http://man7.org/linux/man-pages/man8/tc-netem.8.html.

[5] 360transformations. https://github.com/xmar/360Transformations.

[6] P. C. e. a. Barr, bastien. Multipath tcp: from theory to practice. In *International Conference on Research in Networking*, pages 444–457. 2011, 2011.

[7] V. G. M. Cen S, Cosman P C. End-to-end differentiation of congestion and wireless losses. *IEEE/ACM Transactions on Networking*, 11(5):703–717, 2003.

[8] D. A. e. a. Corbillon X, Simon G. Viewport-adaptive navigable 360-degree video delivery. In *IEEE International Conference on Communications*, pages 1–7. IEEE, 2017.

[9] W. X. e. a. Cui Y, Wang L. Fmtcp:a fountain code-based multipath transmission control protocol. *IEEE/ACM Transactions on Networking*, 23(2):465–478, 2015.

[10] H. S. A. e. a. Duanmu F, Kurdoglu E. Prioritized buffer control in two-tier 360 video streaming. In *Proceedings of the Workshop on Virtual Reality and Augmented Reality Network*, pages 13–18. ACM, 2017.

[11] P. J. Floyd S, Handley M. Equation-based congestion control for unicast applications. In *In Proceedings of SIGCOMM*, pages 43–56. ACM, 2000.

[12] M. C. Graf M, Mueller C. Towards bandwidth efficient adaptive streaming of omnidirectional video over http: Design, implementation, and evaluation. In *ACM on Multimedia Systems Conference*, pages 261–271. ACM, 2017.

[13] L. H. Hwang Y, Obele B O. Multipath transport protocol for heterogeneous multi-homing networks. In *Proceedings of the ACM CoNEXT Student Workshop*, pages 1–2. ACM, 2010.

[14] S. I. The mpeg-dash standard for multimedia streaming over the internet. *IEEE Multimedia*, 18(4):62–67, 2011.

[15] P. Y. e. a. Kwon O C, Go Y. Mpmtp: Multipath multimedia transport protocol using systematic raptor codes over wireless networks. *IEEE Transactions on Mobile Computing*, 14(9):1903–1916, 2015.

[16] C. Y. e. a. Lai Z, Hu Y C. Furion: Engineering high-quality immersive virtual reality on today's mobile devices. In *MobiCom*, pages 409–421. ACM, 2017.

[17] H. Y. S. e. a. Lin S J, Al-Naffouri T Y. Novel polynomial basis with fast fourier transform and its application to reedâĂŞsolomon erasure codes. *IEEE Transactions on Information Theory*, 62(11):6284–6299, 2016.

[18] J. D. Moss and E. R. Muth. Characteristics of headmounted displays and their effects on simulator sickness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(3):308–319, 2011.

[19] H. M. e. a. Petrangeli S, Swaminathan V. An http/2-based adaptive streaming framework for 360Âř virtual reality videos. In *ACM Multimedia Conference*, pages 306–314. ACM, 2017.

[20] R. K. K. Shahryar Afzal, Jiasi Chen. Characterization of 360-degree videos. In *The Workshop on Virtual Reality & Augmented Reality Network*, pages 1–6. ACM, 2017.

[21] K. K. e. a. Sharma V, Kalyanaraman S. A transport protocol to exploit multipath diversity in wireless networks. *IEEE/ACM Transactions on Networking*, 20(4):1024–1039, 2012.

[22] H. W. J. e. a. Sullivan G J, Ohm J. Overview of the high efficiency video coding (hevc) standard. *IEEE Transactions on Circuits & Systems for Video Technology*, 22(12):1649–1668, 2012.

[23] C. N. M. e. a. Wu J, Yuen C. Delay-constrained high definition video transmission in heterogeneous wireless networks with multi-homed terminals. *IEEE Transactions on Mobile Computing*, 15(3):641–655, 2016.

[24] W. M. Wu J, Cheng B. Improving multipath video transmission with raptor codes in heterogeneous wireless networks. *IEEE Transactions on Multimedia*, 20(2):457–472, 2017.

[25] B. Y. e. a. Xie L, Xu Z. 360probdash: Improving qoe of 360 video streaming using tile-based http adaptive streaming. In *IEEE International Conference on Communications*, pages 315–323. ACM, 2017.