

Ατομική Διπλωματική Εργασία

**ΣΥΓΚΡΙΣΗ ΠΡΟΤΥΠΩΝ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΗΝ
ΠΡΟΣΑΡΜΟΣΤΙΚΗ ΜΕΤΑΔΟΣΗ ΒΙΝΤΕΟ**

Ιωάννα Βαλιαντή

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2020

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΣΥΓΚΡΙΣΗ ΠΡΟΤΥΠΩΝ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΗΝ
ΠΡΟΣΑΡΜΟΣΤΙΚΗ ΜΕΤΑΔΟΣΗ ΒΙΝΤΕΟ

Βαλιαντή Ιωάννα

Επιβλέπων Καθηγητής
Δρ. Πατίχης Κωνσταντίνος

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των
απαιτήσεων απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του
Πανεπιστημίου Κύπρου

Μάιος 2020

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή μου κ. Κωνσταντίνο Παττίχη για την ευκαιρία που μου έδωσε να εκπονήσω την παρούσα εργασία και για την εμπιστοσύνη που μου έδειξε. Η υποστήριξή και η καθοδήγησή του διαδραμάτισε πρωταρχικό ρόλο στην επιτυχή ολοκλήρωση αυτής της διπλωματικής εργασίας.

Ιδιαίτερα θα ήθελα να ευχαριστήσω τον Δρ. Αντρέα Παναγίδα για τον πολύτιμο χρόνο και την βοήθεια που μου πρόσφερε καθώς και το ενδιαφέρον που επέδειξε για την εκπλήρωση των στόχων της μελέτης αυτής. Οι εμπειρίες και τις γνώσεις του στον τομέα της επεξεργασίας βίντεο και το πλούσιο υλικό που μου πρόσφερε με βοήθησαν να ολοκληρώσω με επιτυχία την παρούσα εργασία. Του εύχομαι καλή συνέχεια και κάθε επιτυχία.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια και τους φίλους μου για την πολύτιμη στήριξη που μου πρόσφεραν καθ' όλη τη διάρκεια των σπουδών μου στο Πανεπιστήμιο Κύπρου.

Περίληψη

Ο στόχος της παρούσης διπλωματικής εργασίας είναι η σύγκριση υφιστάμενων και αναδυόμενων προτύπων κωδικοποίησης βίντεο για τη χρήση σε συστήματα προσαρμοστικής μετάδοσης βίντεο. Στα πλαίσια της εργασίας, διεξάχθηκαν εκτεταμένα πειράματα με την βοήθεια εργαλείων κωδικοποίησης ανοικτού κώδικα (open-source) εξετάζοντας διαφορετικές απαιτήσεις σε διαθέσιμο εύρος ζώνης (bitrate). Για το σκοπό αυτό, χρησιμοποιήθηκαν συνολικά εικοσιένα βίντεο. Ένα σύνολο από δέκα βίντεο προερχόμενα από τη βάση δεδομένων του Netflix (η οποία δημιουργήθηκε για να εξυπηρετεί αυτού του είδους μελέτες), χωρικής ανάλυσης 768x432, αποτελούμενα από 50 ή 25 πλαίσια ανά δευτερόλεπτο, και ένα σύνολο από τη βάση δεδομένων με σκηνές έκτακτης ανάγκης, χωρικής ανάλυσης 720x576 ή 1440x1080, αποτελούμενα από 25 πλαίσια ανά δευτερόλεπτο. Τα πρότυπα κωδικοποίησης που εξετάστηκαν είναι τα VVC, HM, και x265 προερχόμενα από την σύμπραξη των ISO/IEC MPEG and ITU-T VCEG καθώς και τα AV1 και VP9, προερχόμενα από την νέα εταιρεία Alliance for Open Media (AOM). Για την αξιολόγηση της ποιότητας των βίντεο, χρησιμοποιήθηκαν οι αλγόριθμοι αντικειμενικής αξιολόγησης και σύγκρισης απαιτήσεων εύρους ζώνης, PSNR₆₁₁ και BD-RATE (Bjodegraad Metric), αντίστοιχα. Η πειραματική αξιολόγηση κατέδειξε ότι το σύγχρονο πρότυπο κωδικοποίησης VVC επιτυγχάνει τον καλύτερο λόγο συμπίεσης για το σύνολο των βίντεο που εξετάστηκαν σε σχέση με τα προηγούμενα πρότυπα κωδικοποίησης. Πιο συγκεκριμένα, για την ίδια ποιότητα εικόνας, το πρότυπο VVC μειώνει το απαιτούμενο εύρος ζώνης έως και 26.9% σε σχέση με το αμέσως επόμενο πρότυπο σε απόδοση συμπίεσης, το AV1.

Η διαδικασία που ακολουθήθηκε για την μετάδοση των βίντεο σε πραγματικό χρόνο, αφορούσε την χρήση του προτύπου MPEG-DASH, όπου το βίντεο αναπαράγεται στην οθόνη του παραλήπτη επιλέγοντας από μια γκάμα προ-κωδικοποιημένων βίντεο τα οποία βρίσκονται διαθέσιμα στον εξυπηρετητή του πάροχου του περιεχομένου. Η επιλογή συγκεκριμένου βίντεο γίνεται υπό το πρίσμα του διαθέσιμου εύρους ζώνης, ανακτώντας τμήματα βίντεο, τα οποία μπορεί να έχουν κωδικοποιηθεί με διαφορετικές παραμέτρους συμπίεσης και άρα να έχουν διαφορετικές απαιτήσεις σε εύρος ζώνης και αντίστοιχη ποιότητα. Απώτερος σκοπός είναι τόσο η βελτίωση της ποιότητας όσο και η μείωση του απαιτούμενου εύρους ζώνης κατά την μετάδοση των βίντεο.

Περιεχόμενα

Κεφάλαιο 1	7
Εισαγωγή.....	7
1.1 Εισαγωγή.....	7
1.2 Σκοπός Εργασίας.....	9
1.3 Δομή Εργασίας.....	10
Κεφάλαιο 2	11
Κωδικοποίηση Ψηφιακού Βίντεο	11
2.1 Χαρακτηριστικά του Βίντεο.....	12
2.2 Πλεονασμός	14
2.2.1 Χωρικός πλεονασμός	14
2.2.2 Χρονικός Πλεονασμός	15
2.3 Είδη Κωδικοποίησης.....	15
2.3.1 Ενδοπλαισιακή Κωδικοποίηση (Intra-frame Coding).....	15
2.3.2 Διαπλαισιακή κωδικοποίηση -(Interframe).....	16
2.4 Είδη Πλαισίων κατά την Κωδικοποίηση	17
2.4.1 I-frames	17
2.4.2 P-frames	18
2.4.3 B-Frames	18
2.5 Εφαρμογές Ανοικτού κώδικα και λογισμικό	20
2.5.1 ffmpeg	21
2.5.2 ffserver	22
2.5.3 MP4Box	22
2.6 Κωδικοποιητές	24
2.6.1 HEVC.....	25
2.6.2 x264.....	26
2.6.3 x265.....	26
2.6.4 VP9.....	27
2.6.5 VVC	27
2.6.6 HM	28
2.6.7 AV1.....	28
Κεφάλαιο 3	29

Σύγκριση Προτύπων Κωδικοποίησης	29
3.1 Μεθοδολογία Υλοποίησης	29
3.1.1 Σύνολο Δεδομένων Βίντεο	30
3.1.2 Πειραματική ρύθμιση βίντεο	31
3.2 Αξιολόγηση Αποτελεσμάτων	33
3.2.1 Αλγόριθμος PSNR	34
3.2.2 Ο αλγόριθμος Bjontegaard: BD-Rate Distortion	35
3.3 Αξιολόγηση Κωδικοποιημένων Βίντεο	36
3.4 Συγκεντρωτικά Αποτελέσματα	43
3.5 Σχολιασμός αποτελεσμάτων	45
Κεφάλαιο 4	48
Προσαρμοστική Μετάδοση Βίντεο	48
4.1 Πρωτόκολλα Διάδοσης Βίντεο	49
4.1.1 Progressive Video Streaming	50
4.1.2 Adaptive Video Streaming	51
4.2 MPEG-DASH	53
4.2.1 Περιγραφή MPEG-DASH	54
4.2.2 Media Presentation Description (MPD)	57
4.2.3 Μορφή τμημάτων- Segments	60
4.3 Περιγραφή Συστήματος Προσαρμοστικής Μετάδοσης	61
Κεφάλαιο 5	66
Μέθοδοι Προσαρμοστικής Μετάδοσης Βίντεο	66
5.1 Μέθοδοι Προσαρμοστικής Μετάδοσης Βίντεο	66
5.2 Ασύρματες επικοινωνίες	67
5.2.1 Προσομοίωση Ασύρματων δικτύων	67
5.2.2 Adaptive Controller και TruNET Simulator	68
5.3 Μέθοδος Προσαρμογής του Bitrate	70
Κεφάλαιο 6	75
Συμπεράσματα	75
6.1 Συμπεράσματα για το Σύστημα Προσαρμοστικής Μετάδοσης	75
6.2 Συμπεράσματα για τα Αποτελέσματα από την Σύγκριση Κωδικοποιητών	77
6.3 Μελλοντική Εργασία	78
Βιβλιογραφία	79

Κεφάλαιο 1

Εισαγωγή

1.1 Εισαγωγή	7
1.2 Σκοπός Εργασίας	9
1.3 Δομή Εργασίας	10

1.1 Εισαγωγή

Η μεταφορά αρχείων βίντεο στο διαδίκτυο ξεκίνησε τη δεκαετία του 1990. Ένεκα της ταχείας ανάπτυξης των διαδικτυακών εφαρμογών, οι οποίες αντιμετωπίζουν προβλήματα δικτύου και εύρους ζώνης, επιτακτική καθίσταται η ανάγκη για την ανάπτυξη των αποτελεσματικών προτύπων κωδικοποίησης βίντεο. Μέχρι το 2022 η κυκλοφορία βίντεο on demand, εφαρμογές μετάδοσης βίντεο [1][2]**Error! Reference source not found.** άλλα και διαδικτυακής μάθησης, προσαρμοστική μετάδοση ιατρικού βίντεο, κινητά παιχνίδια και πολλά παιχνίδια σε κινητά τηλέφωνα, προβλέπεται να φτάσουν το 82% της συνολικής κίνησης στο Διαδίκτυο, ποσοστό το οποίο είναι σημαντικά μεγαλύτερο από το 70% που υπήρχε το 2015 [4]. Η καθημερινή ζωή των ανθρώπων περιβάλλεται από συνδεδεμένες συσκευές στο διαδίκτυο, και με τόσες πολλές εφαρμογές αυξάνεται παράλληλα και το Internet traffic με βίντεο. Με την πάροδο των χρόνων, εταιρείες όπως το Youtube αντιμετωπίζουν μεγάλο πρόβλημα video traffic.

Η μετάδοση βίντεο απαιτεί σημαντική υπολογιστική ισχύ, εύρος ζώνης και αρκετή μνήμη, ώστε να παρέχεται βίντεο υψηλής ποιότητας, υπό περιορισμένους πόρους του συστήματος. Όσο οι απαιτήσεις των χρηστών αλλάζουν συνεχώς, η πρόκληση για τις εταιρείες παροχής βίντεο (streaming providers), δεν είναι μόνο να δίνουν το βέλτιστο σύνολο προκαθορισμένων συνδέσεων κωδικοποίησης που ταιριάζουν καλύτερα στο προφίλ του χρήστη, αλλά και να επιλέγουν τη βέλτιστη μορφή κωδικοποίησης βίντεο που

δίνει την καλύτερη δυνατή ποιότητα για ένα συγκεκριμένο χρήστη. Για την αντιμετώπιση αυτής της πρόκλησης αυτής, βασικός πυρήνας της κυκλοφορίας βίντεο αποτελεί η συμπίεσή τους. Οπότεν όλο και μεγαλύτερη είναι η ανάγκη ανάπτυξης αποτελεσματικών αλγορίθμων συμπίεσης, ώστε να ταιριάζουν σε αυτές τις πρωτοφανείς ανάγκες.

Πρόσφατα, παρατηρήθηκε ότι επικρατεί μεγάλη ζήτηση τις συμπίεσης βίντεο σε περιοχές που έχουν χαμηλό εύρος ζώνης, όπως απομακρυσμένες τοποθεσίες και περιοχές που έγιναν καταστροφές **Error! Reference source not found.** Επίσης, επί του παρόντος, τη στιγμή αυτής της Διπλωματικής έχουμε COVID-19 **Error! Reference source not found.** και ως εκ τούτου υπάρχει ένας τεράστιος αριθμός ατόμων που εργάζονται από το σπίτι, οι μαθητές παρακολουθούν διαδικτυακά μαθήματα και πολλά άλλα. Η χρήση του διαδικτύου και η παρακολούθηση βίντεο από σελίδες όπως το Netflix Youtube, αυξήθηκαν ραγδαία, με αποτέλεσμα τα βίντεο που μεταδίδονται συχνά να μην έχουν την βέλτιστη ποιότητα μετάδοσης, ή να σταματούν για ένα μικρό χρονικό διάστημα λόγω του buffering που δημιουργείται. Οι εταιρείες μετάδοσης όμως, σίγουρα δεν θα είναι ευχαριστημένες όταν το βίντεο είναι χαμηλότερης ποιότητας και δεν αφήνει ευχαριστημένο τον χρήστη.

Εκτός από την αυξημένη παρακολούθηση βίντεο σε καθημερινή βάση, λόγω της κανονικής ροής, οι εταιρείες μετάδοσης βίντεο, υπόσχονται ότι παρέχουν πάντα βίντεο υψηλότερης ποιότητας. Για παράδειγμα το Youtube [6] το Netflix, το Facebook κτλ χρησιμοποιούν τις δικές τους ρυθμίσεις κωδικοποίησης, προσπαθώντας να προσαρμόσουν την κωδικοποίηση βίντεο με βάση το περιεχόμενο αλλά και τους περιορισμούς του δικτύου που πιθανόν να υπάρχουν. Αλλά πριν περάσουν στη λύση αυτή, σημαντική είναι η κατανόηση των κωδικοποιητών βίντεο.

Η ποιότητα εμπειρίας του χρήστη (QoE), έχει μεγάλη σημασία όταν έρχεται σε εφαρμογές ροής και Video on Demand, όπου γίνεται η μείωση του εύρους ζώνης. Ένα αποτελεσματικό σύστημα θα ήταν υψηλότερο ή θα είχε αποδεκτή ποιότητα ακόμα και όταν υπάρχει ξαφνική πτώση του εύρους ζώνης για διατήρηση της ποιότητας σε όλο το βίντεο. Τέτοια συστήματα μπορούν να κατασκευαστούν μόνο μετά από κατανόηση της κωδικοποίησης βίντεο με καλύτερη βελτιστοποίηση και έλεγχο.

1.2 Σκοπός Εργασίας

Αυτή η διπλωματική εργασία έχει σαν αρχικό στόχο την κατανόηση τη ορολογίας video streaming, γιατί είναι σημαντική στις μέρες μας και πως επιτυγχάνεται με την χρήση του προτύπου MPEG-DASH.

Η διπλωματική εργασία επικεντρώθηκε σε δύο μέρη. Πρώτα, στόχος της έρευνας ήταν η μελέτη των διαθέσιμων κωδικοποιητών και η σύγκριση της αποδοτικότητας τους μέσω του αλγορίθμου σύγκρισης εύρους ζώνης BD-Rate, και την χρήση του αλγορίθμου αντικειμενικής αξιολόγησης ποιότητας βίντεο PSNR. Σκοπός της μελέτης αυτής είναι η σύγκριση των προτύπων συμπίεσης, όσο αφορά την αποδοτικότητα τους, διερευνώντας τους κωδικοποιητές VP9, HM,x265,AV1 και VVC, σε διαφορετικούς ρυθμούς μετάδοσης έτσι ώστε να δούμε ποιος δίνει το καλύτερο αποτέλεσμα. Υπάρχει ανάγκη για σύγκριση αυτών των προτύπων, λόγω του ότι τα τελευταία χρόνια δημιουργήθηκαν πολλά καινούργια πρότυπα κωδικοποίησης, και λόγω του ότι στην βιβλιογραφία παρατηρούνται αντιφατικά αποτελέσματα. Στόχος ήταν η μελέτη ποιων κωδικοποιητών υπερέχουν για video streaming και συγκεκριμένα για Adaptive Video Streaming.

Στην συνέχεια, επικεντρωθήκαμε στην ανάπτυξη και αξιολόγηση ενός συστήματος, για ποιοτική συμπίεση και μετάδοση βίντεο.. Αυτό επιτεύχθηκε με την δημιουργία ενός συστήματος το οποίο μπορεί να προσαρμόζει το περιεχόμενο που αποστέλλεται, βάσει των μετρήσεων που λαμβάνονται από το δίκτυο. Τα βίντεο στέλνονται από τον αποστολέα με προσαρμοστική μετάδοση σε πραγματικό χρόνο, με την χρήση του προτύπου MPEG-DASH. Μέσω του Adaptive video streaming λύνονται τα προβλήματα ποιότητας και buffering, τα οποία επηρεάζουν τη ροή του βίντεο κατά την διάρκεια προσαρμογής του σε διάφορες συσκευές. Το προτεινόμενο σύστημα μπορεί να βρει εφαρμογή σε οποιοδήποτε είδος βίντεο, τόσο γενικού, όσο και ιατρικού περιεχομένου.

Το σύστημα προσαρμοστικής μετάδοσης βίντεο υλοποιήθηκε μέσω του Dash Player 3.0.1, της Dash Industry Forum, όπου ο Player δίνει πληροφορίες όσο αφορά τα packet-drops, την ποιότητα της εικόνας και άλλες παραμέτρους που αφορούν την ποιότητα της προσαρμοστικής μετάδοσης.

Τέλος, δοκιμάστηκαν οι αλγόριθμοι προσαρμογής με την βοήθεια python scripts και τον προσομοιωτή TruNET, ο οποίος δίνει πληροφορίες όσο αφορά τη διαθεσιμότητα του εύρους. Επαληθεύουμε μέσω της γλώσσας python τον προσαρμοστικό έλεγχο των βίντεο, βάση το bitrate, εξάγουμε στοιχεία που βοηθούν στην αποφυγή του φαινομένου ping pong.

1.3 Δομή Εργασίας

Στην συνέχεια της εργασίας αυτής τα κεφάλαια θα παρουσιαστούν με την εξής δομή:

Το κεφάλαιο 2 περιέχει μια ανάλυση των χαρακτηριστικών των βίντεο καθώς και των τεχνικών κωδικοποίησης.

Το κεφάλαιο 3 περιέχει των προτύπων κωδικοποίησης, την μεθοδολογία που ακολουθήθηκε και τα αποτελέσματα.

Το κεφάλαιο 4 περιέχει μια επεξήγηση των πρωτόκολλων μετάδοσης βίντεο, του τι σημαίνει adaptive video streaming καθώς και λεπτομερή ανάλυση του Πρωτοκόλλου MPEG-DASH. Ακολουθεί η επεξήγηση της ανάπτυξη του συστήματος προσαρμοστικής μετάδοσης που δημιουργήθηκε.

Το κεφάλαιο 5 αναφέρεται στους μεθόδους προσαρμοστικής μετάδοσης βίντεο και συγκεκριμένα παρουσιάζεται μια προσομοίωση σε python που αφορά τη προσαρμογή του bitrate στο διαθέσιμο εύρος ζώνης του δικτύου.

Το κεφάλαιο 6 παρουσιάζει τα συμπεράσματά που προκύπτουν, μέσα από τους βασικούς στόχους που είχε αυτή η διπλωματική εργασία.

Κεφάλαιο 2

Κωδικοποίηση Ψηφιακού Βίντεο

2.1 Χαρακτηριστικά του Βίντεο	12
2.2 Πλεονασμός	14
2.2.1 Χωρικός πλεονασμός	14
2.2.3 Χρονικός Πλεονασμός / Υποκειμενικός Πλεονασμός	15
2.3 Είδη Κωδικοποίησης	15
2.3.1 Ενδοπλαισιακή Κωδικοποίηση (Intra-frame Coding)	15
2.3.2 Διαπλαισιακή κωδικοποίηση -(Interframe)	16
2.4 Είδη Πλαισίων κατά την Κωδικοποίηση	17
2.4.1 I-frames	17
2.4.2 P-frames	18
2.4.3 B-Frames	18
2.5 Εφαρμογές Ανοικτού κώδικα και λογισμικό	20
2.5.1 ffmpeg	21
2.5.2 ffserver	22
2.5.3 MP4Box	22
2.6 Κωδικοποιητές	24
2.6.1 HEVC	25
2.6.2 x264	26
2.6.3 x265	26
2.6.4 VP9	27
2.6.5 VVC	27
2.6.6 HM	28
2.6.7 AV1	28

Σε αυτή την ενότητα αρχικά επεξηγούνται τα βασικά χαρακτηριστικά του βίντεο, τα οποία επηρεάζουν την ποιότητα της κωδικοποίησης. Στην συνέχεια αναφέρονται τα

βασικά είδη κωδικοποίησης που υπάρχουν καθώς και τα είδη των πλαισίων που δημιουργούνται κατά την διάρκεια της συμπίεσης. Αναλύονται οι εφαρμογές ανοικτού κώδικα και τα πρότυπα κωδικοποίησης που χρησιμοποιήθηκαν στην παρούσα μελέτη.

2.1 Χαρακτηριστικά του Βίντεο

Η αντίληψη της κίνησης σε ένα βίντεο βασίζεται σε δύο οπτικές ψευδαισθήσεις, το phi και το beta φαινόμενο [7]. Το phi φαινόμενο, βασίζεται στην οπτική ψευδαίσθηση που έχει ο χρήστης ότι ενώ αντιλαμβάνεται την κίνηση σε μια ακολουθία, το οπτικό σύστημα «συμπληρώνει» τις πληροφορίες που λείπουν. Το beta φαινόμενο είναι η ψευδαίσθηση της κίνησης που δημιουργείται όταν τα ερεθίσματα αλλάζουν θέση σε μια ακολουθία εικόνων. Το ανθρώπινο μάτι αντί να αντιλαμβάνεται τις εικόνες σε μια σειρά, έχει την αίσθηση ότι υπάρχει κίνηση. Η εναλλαγή διαδοχικών εικόνων (καρέ) προκαλεί αυτή την αίσθηση κίνησης, όμως αυτό συμβαίνει μόνο όταν η εναλλαγή είναι μεγαλύτερη από 25 φορές το δευτερόλεπτο[9]. Ένα βίντεο αποτελείται από διαδοχικές εικόνες, οι οποίες προβάλλονται γρήγορα, έτσι ώστε να δίνεται η αίσθηση ότι τα αντικείμενα κινούνται. Οι διαδοχικές εικόνες ονομάζονται πλαίσια ή καρέ (frames). Εκτός από την κατάσταση του δικτύου και την προσπάθεια βελτίωσής του, με στόχο το καλύτερο δυνατό αποτέλεσμα (QoE), έμφαση πρέπει να δίνεται και στο περιεχόμενο που πρόκειται να μεταδοθεί και τον τρόπο κωδικοποίησης του. Βασικά χαρακτηριστικά των βίντεο τα οποία επηρεάζουν την απόδοση της κωδικοποίησης, θεωρούνται το περιεχόμενο, τα Frames per second (fps), και ανάλυση (video resolution) [9].

Frames per second (fps)

Είναι ο αριθμός των πλαισίων που προβάλλονται ανά δευτερόλεπτο κατά την αναπαραγωγή ενός βίντεο. Σε αυτό τον παράγοντα βασίζεται η ομαλή εμφάνιση του βίντεο, δηλαδή πρέπει να είναι μεγάλος ο αριθμός (π.χ. 25 fps/30 fps), ώστε να μην υπάρχουν απότομες εναλλαγές στις εικόνες του βίντεο. Το ανθρώπινο μάτι έχει την ιδιότητα να αντιλαμβάνεται τη ροή διαδοχικών εικόνων σαν να είναι συνεχόμενη, στην περίπτωση που το frame rate είναι μεγαλύτερο από 16 έως 18 fps.

Ανάλυση Βίντεο (video resolution)

Η ανάλυση βίντεο, γνωστή και ως χωρική ανάλυση (spatial resolution) είναι το γινόμενο της οριζόντιας με την κάθετη διάστασης της εικόνας. Εκφράζει το πλήθος των pixels (εικονοστοιχεία) που αποτελούν ένα πλαίσιο. Για παράδειγμα, βίντεο με ανάλυση 640×480 σημαίνει ότι οι εικόνες που αποτελούν το βίντεο έχουν διάσταση 640 pixels στον οριζόντιο άξονα και 480 pixels στον κάθετο άξονα. Γενικά, μεγαλύτερη ανάλυση ισοδυναμεί σε εικόνα μεγαλύτερης διάστασης. Στους σύγχρονους υπολογιστές η ανάλυση του βίντεο είναι συχνά αρκετά υψηλή και φτάνει στα 1920×1080. Η ανάλυση στους σύγχρονους υπολογιστές είναι γνωστή ως Full HD (Full High Definition) και ουσιαστικά αποτελείται από 1080 οριζόντιες γραμμές. Συνήθως [10] αναφέρεται είτε ως Full HD είτε ως 1080p. Συνήθεις αναλύσεις βίντεο που προβάλλονται μέσω του διαδικτύου, π.χ. στο YouTube, είναι οι 360p (480×360), 480p (640×480) και 720p (1280×720). Η ανάλυση 720p είναι αρκετά υψηλή για διαδικτυακό βίντεο και μπορεί να υποστηριχθεί με τις τρέχουσες ταχύτητες στο διαδίκτυο. Η ανάλυση Full HD απαιτεί υψηλές ταχύτητες μεταφοράς δεδομένων στο διαδίκτυο και σε αρκετές περιπτώσεις δεν χρησιμοποιείται σε απευθείας μετάδοση. Με την εξέλιξη της τεχνολογίας έχει εμφανιστεί και ο όρος UHD (Ultra High Definition). Συναντάται, επίσης, με την ορολογία 4K και ουσιαστικά αναφέρεται σε όλες τις αναλύσεις που είναι υψηλότερες από το Full HD. Αναφέρεται ως ανάλυση 2160p (3840×2160) και μέχρι σήμερα φτάνει το 8K UHD, δηλαδή ανάλυση 7680×4320 [9].

Περιεχόμενο Βίντεο

Το επίπεδο συμπίεσης του βίντεο εξαρτάται από το περιεχόμενο του βίντεο. Για παράδειγμα, όταν το βίντεο έχει περιεχόμενο με γρήγορη κίνηση χρειάζεται υψηλότερο frames per second από ότι απαιτούν τα βίντεο στατικού περιεχομένου, (παρουσιάσεις slideshow, video conference). Η χρονική διάρκεια (duration) είναι ο χρόνος που διαρκεί η προβολή του βίντεο και εκφράζεται σε δευτερόλεπτα.

Bitrate

Ο ρυθμός μετάδοσης Bit Rate, δείχνει τον αριθμό των bit ανά δευτερόλεπτο (bps-bits per second). Ο ρυθμός μετάδοσης εξαρτάται από την μέθοδο κωδικοποίησης που χρησιμοποιείται, και σχετίζεται με τα δεδομένα που περιέχονται σε ένα δευτερόλεπτο μιας ροής βίντεο. Όσο μεγαλύτερος είναι ο ρυθμός μετάδοσης τόσο καλύτερη είναι η

ποιότητα του βίντεο, καθώς μπορούν να μεταδοθούν περισσότερα δεδομένα, άρα βίντεο μεγαλύτερης ανάλυσης. Ο ρυθμός μετάδοσης bit μπορεί να είτε σταθερός (CBR – Constant Bit Rate), είτε μεταβλητός (VBR – Variable Bit Rate). Στον μεταβλητό bitrate επιτυγχάνεται καλύτερη ποιότητα βίντεο, καθώς σε σκηνές με γρήγορη κίνηση, το bitrate είναι υψηλότερο από ότι σε σκηνές με αργή κίνηση.

2.2 Πλεονασμός

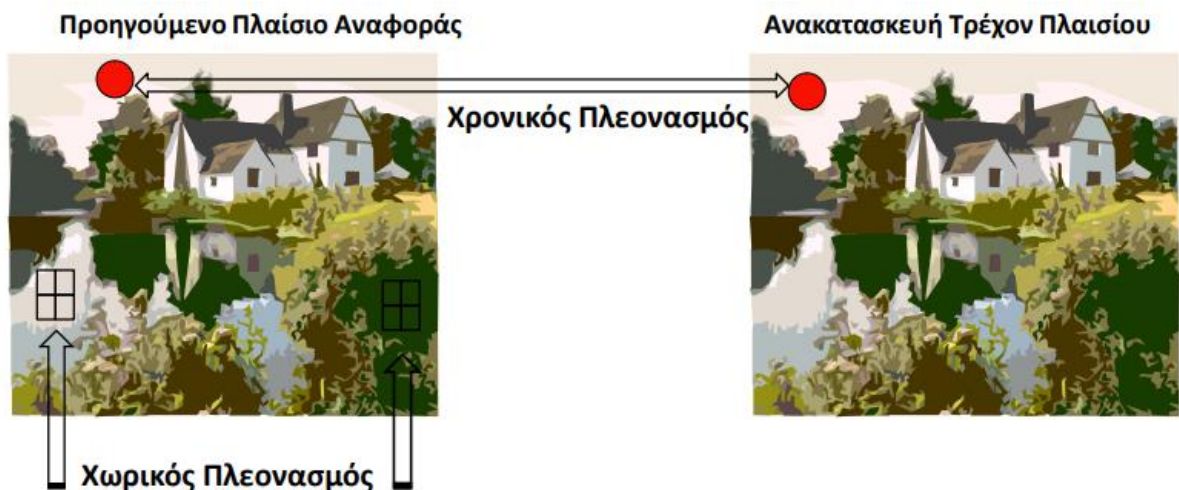
Η ποσότητα της πληροφορίας ενός βίντεο εξαρτάται από τον αριθμό των πλαισίων ανά δευτερόλεπτο που απαιτούνται για την αναπαραγωγή της και από την ποιότητα της εικόνας. Ένα βίντεο χρειάζεται μεγάλο όγκο δεδομένων για να αποσταλεί ή να αποθηκευτεί. Για να γίνει πιο εύκολο αυτό είναι απαραίτητο να γίνεται συμπίεση των δεδομένων. Η συμπίεση των δεδομένων βασίζεται στον πλεονασμό. Με τον όρο αυτό η πληροφορία είτε μπορεί να παραληφθεί, είτε να συμπιεστεί χωρίς να προκαλεί σημαντική επίδραση στο τελικό αποτέλεσμα. Με την χρήση του πλεονασμού μπορεί να μειωθεί ο όγκος των δεδομένων που μεταφέρονται ή αποθηκεύονται. Ο πλεονασμός παρουσιάζεται όταν δημιουργούνται τα δεδομένα και κατηγοριοποιείται σε χωρικό πλεονασμό (Spatial Redundancy) και σε χρονικό πλεονασμό (Temporal Redundancy)[11].

2.2.1 Χωρικός πλεονασμός

Προκύπτει γιατί τα γειτονικά εικονοστοιχεία (pixel) που εμπεριέχονται σε ένα frame συχνά είναι όμοια μεταξύ τους. Οι τιμές των εικονοστοιχείων μπορούν να υπολογιστούν με μεγάλη ακρίβεια χρησιμοποιώντας τα γειτονικά τους. Με αυτό τον τρόπο εκμεταλλεύεται τις ιδιότητες της ανθρώπινης όρασης, αφού μικρές αλλαγές στην αλλοίωση μιας εικόνας ή βίντεο δεν γίνονται αντιληπτά και αυτό γιατί το μάτι είναι πιο ευαίσθητο στην φωτεινότητα μιας εικόνας και στα χρώματα. Για την παρουσίαση των χρωμάτων και της πλεονάζουσας πληροφορίας μπορεί να χρησιμοποιείται λιγότερο εύρος ζώνης από αυτό που είναι διαθέσιμο, χωρίς να το αντιλαμβάνεται ο θεατής. Για παράδειγμα κατά την διάρκεια ενός δελτίου ειδήσεων, μόνο ένα κομμάτι από το πρόσωπο του παρουσιαστή αλλάζει, τα χείλη, τα μάτια και τα βλέφαρα, το φόντο της εικόνας δεν αλλάζει. Με αυτό τον τρόπο ένα τμήμα της εικόνας μπορεί να υπολογιστεί από μια προηγούμενη, προσθέτοντας σε αυτή τις διαφορές που έχουν προκύψει, αποφεύγοντας την κωδικοποίηση της πλεονάζουσας πληροφορία σε κάθε πλαίσιο.

2.2.2 Χρονικός Πλεονασμός

Ο χρονικός ή αλλιώς υποκειμενικός πλεονασμός (Temporal Redundancy) προκύπτει όταν ένα κομμάτι της παρόν εικόνας σχετίζεται με ένα τμήμα της εικόνας του προηγούμενου πλαισίου. Εκμεταλλεύοντας το ότι διαδοχικά πλαίσια παρουσιάζουν μεγάλες ομοιότητες, υπολογίζονται κάποια κομμάτια της εικόνας από την προηγούμενη και στη συνέχεια να προθέτονται σε αυτή οι διαφορές που έχουν προκύψει, με αυτό τον τρόπο αποφεύγεται η κωδικοποίηση ολόκληρης της εικόνας που έχει πλεονάζουσα πληροφορία.



Σχήμα 2.2: Χωρικός και Χρονικός Πλεονασμός[12]

2.3 Είδη Κωδικοποίησης

Ένα βίντεο αποτελείται από μια ακολουθία από πλαίσια (frames), τα οποία προβάλλονται συνεχόμενα το ένα μετά το άλλο. Για να κωδικοποιηθεί ένα βίντεο, αρκεί να κωδικοποιηθούν τα επιμέρους πλαίσιά του.

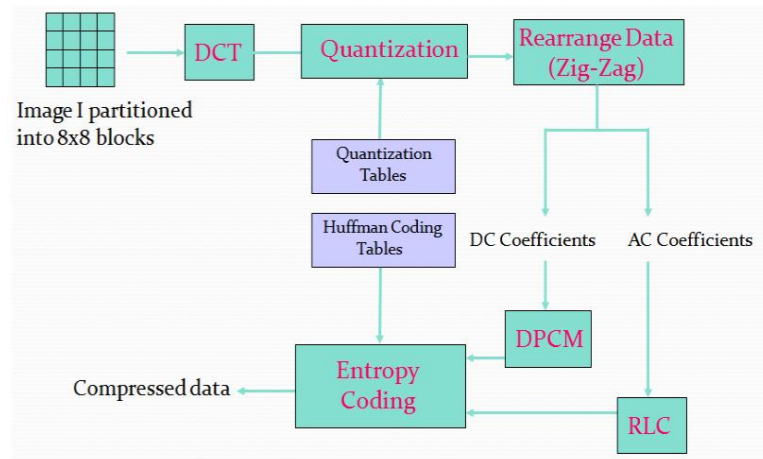
2.3.1 Ενδοπλαισιακή Κωδικοποίηση (Intra-frame Coding)

Η Ενδοπλαισιακή Κωδικοποίηση αποτελείται μόνο από I-frames. Κάθε πλαίσιο κωδικοποιείται ανεξάρτητα από τα υπόλοιπα, χωρίς να χρησιμοποιεί άλλα πλαίσια του βίντεο. Επιτυγχάνεται μείωση του χωρικού πλεονασμού, αφού το τελικό σήμα αποτελείται από μια σειρά από διακριτές και ακίνητες εικόνες.

Από την μία πλευρά, για την μετάδοση του βίντεο απαιτείται μεγάλο εύρος ζώνης. Αυτό συμβαίνει γιατί κάθε πλαίσιο περιέχει όλη την πληροφορία που χρειάζεται για να κωδικοποιηθεί σαν μια ανεξάρτητη εικόνα και να αναπαρασταθεί μετά την

αποκωδικοποίηση. Από την άλλη πλευρά, μέσω της μεθόδου αυτής δεν επηρεάζονται άλλα πλαίσια στην περίπτωση που υπάρχει λάθος στην μετάδοση κάποιου πλαισίου.

Η κωδικοποίηση του βίντεο επιτυγχάνεται μέσω της κωδικοποίησης των επιμέρους πλαισίων που το αποτελούν, δηλαδή διαχωρίζονται σε macroblocks, γίνεται μετασχηματισμός DCT, κβαντοποίηση και κωδικοποίηση εντροπίας, όπως φαίνεται στο σχήμα 2.3.



Σχήμα 2.3: Διάγραμμα Ροής JPEG[10]

2.3.2 Διαπλαισιακή κωδικοποίηση -(Interframe)

Στην διαπλαισιακή κωδικοποίηση το βίντεο αποτελείται και από P και B-frames, χρησιμοποιώντας αντιστάθμιση κίνησης. Το εύρος ζώνης είναι μικρότερο σε σχέση με την ενδοπλαισιακή κωδικοποίηση. Σε περίπτωση σφάλματος, όμως, σε πλαίσιο αναφοράς, το σφάλμα μεταφέρεται στο πλαίσιο που πρόκειται να προβλεφθεί, αφού κωδικοποιείται μόνο το πρώτο πλαίσιο και για τα υπόλοιπα υπολογίζονται μόνο οι διαφορές με το πρώτο. Για τον υπολογισμό του i-οστού πλαισίου, χρησιμοποιείται το πρώτο πλαίσιο και η διαφορά του από το i-οστό πλαίσιο, η διαφορά κωδικοποιείται ως ανεξάρτητο πλαίσιο (ενδοπλαισιακή κωδικοποίηση), γιατί αποτελεί πλαίσιο αναφοράς. Τα πλαίσια αξιοποιούν τον χωρικό πλεονασμό, δηλαδή όταν σε ένα πλαίσιο υπάρχουν ίδια στοιχεία με ίδιες τιμές, αλλά αξιοποιούν και τον χρονικό πλεονασμό, δηλαδή ότι δύο πλαίσια έχουν ίδιες τιμές σε ίδιες θέσεις και έτσι έχουμε καλύτερη απόδοση συμπίεσης.

2.4 Είδη Πλαισίων κατά την Κωδικοποίηση

Όταν το βίντεο συμπιέζεται, αποτελείται από πλαίσια (Frames), τα οποία διακρίνονται σε τρεις βασικές κατηγορίες I-frames, P-frames, B-frames. με βάση της ιδιότητες και τα χαρακτηριστικά τους [13].

2.4.1 I-frames

Τα I-frames αποτελούν τα πιο βασικά πλαίσια (frame) και λειτουργούν αυτόνομα, ονομάζονται και πλαίσια αναφοράς. Κάθε πλαίσιο / block λαμβάνεται ως ξεχωριστή οντότητα και κωδικοποιείται ξεχωριστά και ανεξάρτητα, χωρίς να χρειάζεται να γίνει αναφορά σε κάποια ανεξάρτητα ή επόμενα πλαίσια. Ανεξάρτητα λειτουργούν και κατά την αποκωδικοποίηση, αφού περιέχουν αναφορές μόνο στον εαυτό τους και δεν χρειάζονται αναφορές σε άλλα πλαίσια. Τα I-frames χρησιμοποιούν Intra frame coding (ενδοπλαισιακή κωδικοποίηση). Ο αριθμός των πλαισίων μεταξύ διαδοχικών I-frames, είναι γνωστός ως group of pictures (GOP). Η δομή ενός τέτοιου συνόλου εικόνων ξεκινά από ένα I-frame και στην συνέχεια ακολουθούν τα P και B-frames. Με αυτό τον τρόπο επιτυγχάνεται τυχαία προσπέλαση της ακολουθίας και λειτουργίες όπως reverse και fast forward.

Αφού λειτουργούν σαν σημεία αναφοράς, πρέπει να μεταδίδονται ανά τακτικά χρονικά διαστήματα σε μια εξερχόμενη ροή. Είναι σημαντικό, καθώς αν ένα I-frame είναι διεφθαρμένο, τότε αυτό θα προκαλέσει πρόβλημα, στα πλαίσια τα οποία κάνουν αναφορά σε αυτό για την πρόβλεψή τους, ως επίσης και για να αποφεύγεται η διάδοση σφαλμάτων που μπορεί να δημιουργηθούν από τα P-frames. Για παράδειγμα, αν η συχνότητα είναι 30Hz, τουλάχιστον κάθε 15 πλαίσια πρέπει να υπάρχει ένα I-frame, δηλαδή δύο φορές το δευτερόλεπτο.

Η κωδικοποίηση των πλαισίων σχετικά με τα στοιχεία Y,U,V, γίνεται ανεξάρτητα και βασίζεται στην λογική του αλγορίθμου JPEG. Κατά την κωδικοποίηση JPEG, η εικόνα χωρίζεται σε macroblocks και για κάθε block ξεχωριστά εφαρμόζεται DCT, κβαντοποίηση, zig zag, run-length encoding και Huffman encoding.

Όσο αφορά την αποκωδικοποίηση, τα I-frames περιέχουν όλη την απαραίτητα πληροφορία που χρειάζεται ώστε το πλαίσιο να μπορεί να αναπαρασταθεί αυτόνομα, χωρίς να εξαρτάται από τα υπόλοιπα πλαίσια. Μόλις το πλαίσιο ληφθεί, αποκωδικοποιούνται οι συμπιεσμένες πληροφορίες, ώστε να ξαναδημιουργηθεί το

αρχικό πλαίσιο. Κατά συνέπεια, καταλαμβάνεται περισσότερος χώρος αποθήκευσης και χαμηλό επίπεδο συμπίεσης.

2.4.2 P-frames

Τα P-frames ονομάζονται προβλεπόμενα λόγω του ότι υπολογίζονται με πρόβλεψη, χρησιμοποιώντας δηλαδή το πιο κοντινό προηγούμενο I-frame ή P-frame. Ο αριθμός των πλαισίων μεταξύ ενός P-frame και του αμέσως προηγούμενου I-frame ή P-frame ονομάζεται «prediction span». Δημιουργούνται και συμπιέζονται μετά τα I-frame και είναι γνωστά και ως «Anchor frame».

Έχουν πολύ μικρότερο μέγεθος από τα I-frames, καθώς έχουν μεγαλύτερο ποσοστό συμπίεσης, λόγω του ότι η κωδικοποίηση βασίζεται στην αντιστάθμιση της κίνησης (motion compensation), που προβλέπει τη νέα θέση όποιων macroblock έχουν μετακινηθεί. Συνεπώς, δεν χρειάζεται να κωδικοποιηθεί ολόκληρη η εικόνα, αλλά μόνο η πληροφορία που την σχετίζει με την εικόνα αναφοράς και ονομάζεται διάνυσμα κίνησης (motion vector), αξιολογείται με αυτό τον τρόπο ο χρονικός πλεονασμός.

Για την δημιουργία του τελικού αποκωδικοποιημένου πλαισίου, χρησιμοποιείται η αποκωδικοποιημένη πληροφορία με το αποκωδικοποιημένο προηγούμενο I-frame ή P-frame.

Είναι πιο συμπεσμένα από τα I-frames, όμως μειονέκτημα θεωρείται το ότι δεν έχουν την ίδια ακρίβεια και το γεγονός ότι είναι ευαίσθητα σε λάθη μετάδοσης, αφού εξαρτώνται από προηγούμενα πλαίσια. Μπορούν να λειτουργήσουν και σαν σημεία αναφοράς για επόμενα πλαίσια, λόγω του ότι λαμβάνουν μέρος στην δημιουργία και διάδοση σφαλμάτων, ενός GOP, εξαιτίας των δομημένων συσχετίσεων στο οποίο βασίζεται η κωδικοποίηση τους, αφού η διαδικασία της πρόβλεψης της κίνησης (motion estimation) δεν είναι 100% ακριβής. Μέσα από την διαδικασία κωδικοποίησης, συμπεραίνουμε ότι ένα P-frame αποτελείται από τα διανύσματα κίνησης, τα Intra-coded blocks, τα οποία δεν συσχετίστηκαν με κάποιο προηγούμενο πλαίσιο, και την διαφορά του εκτιμώμενου πλαισίου από το πραγματικό.

2.4.3 B-Frames

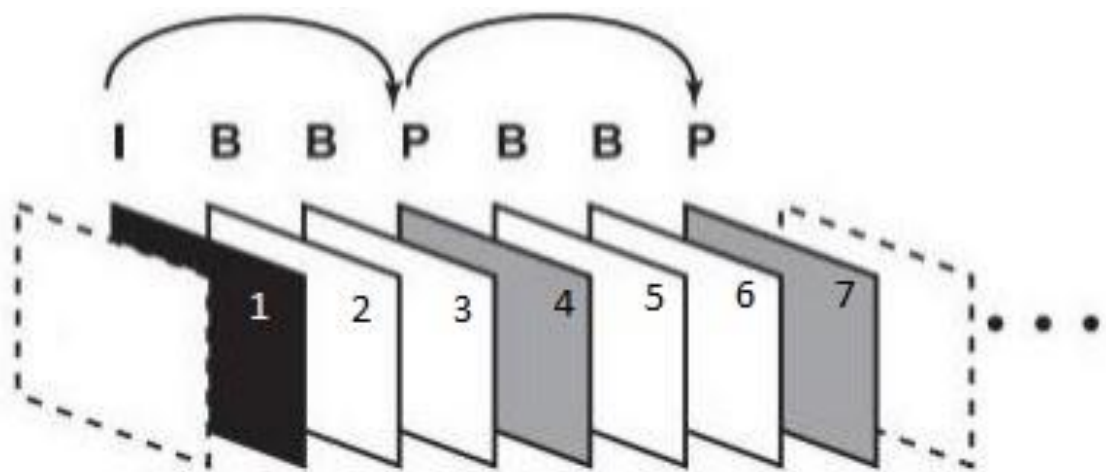
Τα B-frames, ονομάζονται αμφίδρομα λόγω του ότι μπορούν να προβλέψουν τα προηγούμενα ή τα επόμενα I-frame και P-frame. Παίρνουν τον μέσο όρο σε επίπεδο

macroblock των προηγούμενων πλαισίων και γι' αυτό μειώνουν σε μεγάλο βαθμό τα σφάλματα. Τα B-frames δεν λειτουργούν σαν σημεία αναφοράς και έτσι δεν λαμβάνουν μέρος στην διάδοση σφαλμάτων σε ένα GOP, καθώς η εκτίμησή τους περιορίζεται μόνο μεταξύ δυο διαδοχικών πλαισίων.

Είναι τα πλαίσια με το μικρότερο μέγεθος και εκμεταλλεύονται σε μεγάλο βαθμό τον χρονικό πλεονασμό μεταξύ δύο P-frames. Είναι περισσότερο συμπιεσμένα από τα P-frames και η διαφορά τους από αυτά είναι ότι μπορούν να προβλέψουν και επόμενα πλαίσια. Με αυτό τον τρόπο επιτυγχάνεται καλύτερη εκτίμηση της κίνησης, όμως αυξάνεται η καθυστέρηση της κωδικοποίησης άλλων πλαισίων, αφού δεν γίνεται αναφορά σε αυτά. Σε περίπτωση όπου το B-frame βρίσκεται τελευταίο σε ένα GOP, πρέπει να χρησιμοποιηθεί το προηγούμενο P-frame και το επόμενο I-frame για την εκτίμησή του. Άρα για να μπορεί να αποκωδικοποιηθεί και να προβληθεί ένα B-frame είναι απαραίτητο να αποκωδικοποιηθεί πρώτα το επόμενο I-frame ή P-frame που ακολουθεί ένα B-frame, έτσι ώστε να μπορεί να αποκωδικοποιηθεί και να προβληθεί. Η διαδικασία αυτή ωστόσο θεωρείται πολύπλοκη υπολογιστικά, καθώς απαιτεί μεγάλα data buffers που προκαλούν μεγάλη καθυστέρηση κατά την κωδικοποίηση και την αποκωδικοποίηση, γι' αυτό και πολλές φορές τα B-frames δεν είναι επιθυμητά από τους κατασκευαστές.

Όσο αφορά την αποκωδικοποίηση, για να δημιουργηθεί το τελικό αποκωδικοποιημένο πλαίσιο, πρέπει πρώτα να αποκωδικοποιηθεί η πληροφορία που φτάνει και μετά να αποκωδικοποιηθεί και το αμέσως προηγούμενο και επόμενο I-frame ή P-frame. Αλλάζοντας την σειρά κωδικοποίησης και μετάδοσης των κωδικοποιημένων πλαισίων, τότε χρόνος που απαιτείται για την αποκωδικοποίηση κάθε B-frame μπορεί να μειωθεί. Έτσι όταν λαμβάνεται ένα B-frame, τότε είναι διαθέσιμα το αμέσως προηγούμενο και επόμενο I-frame ή P-frame. Με αυτό τον τρόπο είναι διαθέσιμες οι αποκωδικοποιημένες πληροφορίες των δύο πλαισίων στα οποία γίνεται αναφορά.

Η ποιότητα του συμπιεσμένου βίντεο καθορίζεται σε μεγάλο βαθμό από το πόσο συχνά εμφανίζονται τα πλαίσια αναφοράς, ανάλογα με το περιεχόμενο του βίντεο επιλέγεται και ο αριθμός αυτός GOP. Στις πλείστες περιπτώσεις το πρώτο πλαίσιο είναι τύπου I και τα υπόλοιπα απεικονίζουν τις διαφορές με το πρώτο τύπου P,B. Η σειρά με την οποία εμφανίζονται τα πλαίσια διαφέρει από την σειρά με την οποία κωδικοποιούνται. Για παράδειγμα, για την πρόβλεψη ενός B-frame, πρέπει πρώτα να κωδικοποιηθεί το αμέσως προηγούμενο και επόμενο P-frame, όπου αναφέρεται. Όταν όμως το βίντεο αναπαραχθεί το B-frame θα βρίσκεται πριν από το δεύτερο P-frame που χρησιμοποιήθηκε. Ο χρονικός πλεονασμός μειώνεται από τις συσχετίσεις που δημιουργούνται μεταξύ των διαφόρων τύπων πλαισίων, όπως παρουσιάζεται στην στο σχήμα 2.4.



Σχήμα 2.4: Ένα τυπικό GOP, που παρουσιάζει τη χρονική ακολουθία μεταξύ I, P και B πλαισίων. Σειρά Κωδικοποίησης πλαισίων: 1,4,2,3,7,5,6

2.5 Εφαρμογές Ανοικτού κώδικα και λογισμικό

Λογισμικό Ανοικτού Πηγαίου Κώδικα (Open Source Software -OSS) ονομάζεται το λογισμικό, του οποίου εκτός από το εκτελέσιμο αρχείο διατίθεται απευθείας και ο πηγαίος κώδικας του, ο οποίος μπορεί να μεταγλωττιστεί. Αποτελεί μια αποκεντριοποιημένη ανάπτυξη όπου συνήθως βασίζεται σε παγκόσμια κλίμακα μέσω κοινόχρηστων repositories και οι Χρήστες μπορεί να εισηγηθούν αλλαγές οι οποίες υλοποιούνται δωρεάν & γρήγορα, ειδικά εάν συντρέχουν λόγοι ασφάλειας. Ο κώδικας ενός λογισμικού πολλές φορές είναι σημαντικό να διατίθεται για περαιτέρω μελέτη κατανόηση και βελτιστοποίηση του προγράμματος. Προφανώς, για να επιτευχθεί αυτό τίθενται θέματα πνευματικών δικαιωμάτων, τα οποία καλύπτονται από ειδικές άδειες. Αρχικά υπάρχει το Free OSS (FOSS), Λογισμικό Ανοικτού Πηγαίου Κώδικα το οποίο

είναι δωρεάν, δηλαδή επιτρέπει στους χρήστες να το διανέμουν ως μέρος της υλοποίησης τους ή και να το επεκτείνουν. Το ζητούμενο του FOSS είναι πως θα διασφαλίσεις ότι η ανοικτή γνώση (ένα κομμάτι ανοικτού πηγαίου κώδικα) θα συνεχίσει να παραμένει ανοικτή γνώση. Οι άδειες χρήσης FOSS χωρίζονται σε τρεις κατηγορίες ως εξής:

- Strong Copyleft: όπως για παράδειγμα το GPL, όπου το λογισμικό διατίθεται δωρεάν για χρήση αλλά οι επεκτάσεις επιβάλλεται να παραμείνουν ανοικτές, δηλαδή το καινούργιο παραγόμενο λογισμικό πρέπει να έχει την ίδια άδεια με το λογισμικό που χρησιμοποιεί αρχικά. Το παραγόμενο εκτελέσιμο αρχείο μπορεί να διατίθεται δωρεάν ή με κάποια τιμή, όμως ο παραγόμενος πηγαίος κώδικας πρέπει να παραμείνει ελεύθερος και δωρεάν.
- Weak Copyleft: Ενδιάμεσο, όπως Lesser GPL, όπου η χρήση της βιβλιοθήκης μπορεί να γίνει δωρεάν αλλά υπάρχει περίπτωση να γίνει link με κλειστό λογισμικό. Δηλαδή σε αντίθεση με το GPL, αν κάποιος χρησιμοποιήσει ένα ήδη παραγόμενο λογισμικό το οποίο προήλθε από GPL, τότε δεν είναι υποχρεωμένος να διαθέσει τον κώδικα του εκτελέσιμο ως FOSS, αλλά μπορεί να το κρατήσει ως κλειστό λογισμικό.
- Permissive, Ανεκτική: όπως MIT και BSD, δηλαδή το λογισμικό μπορεί να χρησιμοποιηθεί και να τροποποιηθεί δωρεάν και οι επεκτάσεις μπορεί να παραμείνουν κλειστές [14].

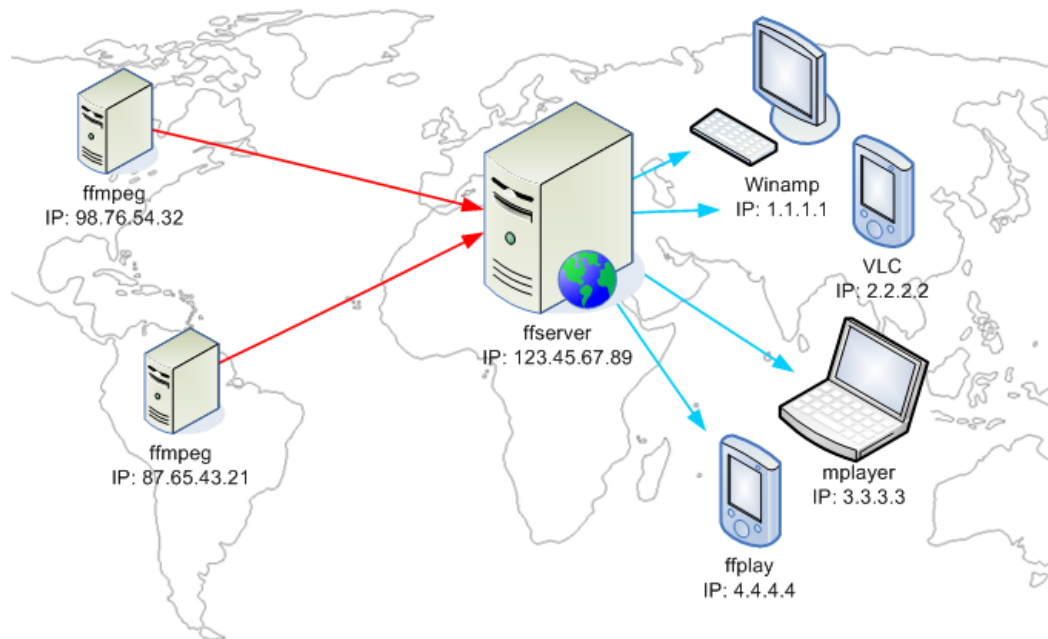
2.5.1 ffmpeg

Το ffmpeg είναι ένα δωρεάν λογισμικό ανοικτού κώδικα που αποτελείται από ποικίλες βιβλιοθήκες και προγράμματα για την διαχείριση βίντεο και ήχου, μέσω της γραμμής εντολών (command line). Ο όρος προήλθε από ένα συνδυασμό του προτύπου MPEG με το ff, που είναι τα αρχικά του fast forward. Ανάλογα με τις επιλογές που θα χρησιμοποιηθούν είναι διαθέσιμο κάτω από το GNU Lesser General Public License ή το GNU General Public License. Οι βιβλιοθήκες του ffmpeg χρησιμοποιούνται στα λογισμικά των media players, όπως το VLC. Μπορεί επίσης να καταγράψει και να κωδικοποιήσει σε πραγματικό χρόνο, από διάφορες πηγές υλικού, όπως για παράδειγμα TV capture card. Τα εργαλεία του ffmpeg, όπως, ο ffmpeg, ο ffplay και ο ffprobe, χρησιμοποιούνται για την επιτυχή κωδικοποίηση, αποκωδικοποίηση, αναπαραγωγή, streaming και φιλτράρισμα αρχείων ήχου και βίντεο. Το ffplay είναι ένα απλό πρόγραμμα αναπαραγωγής πολυμέσων (media player), το ffprobe είναι ένα εργαλείο που μέσω της

γραμμής εντολών παρουσιάζει πληροφορίες των αρχείων, ενώ ο ffmpeg περιγράφεται πιο κάτω [15].

2.5.2 ffmpeg

Ο ffmpeg είναι ένα εργαλείο που παρέχεται μέσω του ffmpeg, είναι ένας server συνεχούς ροής (broadcasting) όπου μεταφέρεται το περιεχόμενο του βίντεο μέσω του διαδικτύου. Ο ffmpeg χρησιμοποιεί ένα configuration file, στο οποίο καθορίζονται τα port στα οποία ο server θα στείλει τα κωδικοποιημένα βίντεο. Όπως φαίνεται στο πιο κάτω σχήμα, για κάθε κωδικοποίηση μπορούμε να έχουμε πολλές ροές εξόδου.



Σχήμα 2.5: Διάφορες πηγές εισόδου (ffmpeg), μπορούν να χρησιμοποιηθούν για την «τροφοδοσία» του ffmpeg, ο οποίος με την σειρά του διανέμει το υλικό σε πολλούς πελάτες για προβολή [16].

2.5.3 MP4Box

MP4 Box είναι ένας συσκευαστής πολυμέσων MP4 που διατίθεται από το ανοικτού κώδικα GPAC (Project on Advanced Content). Χρησιμοποιείται για την αναπαραγωγή και λήψη αρχείων πολυμέσων όπως AVI, MPG και κυρίως σε αρχεία πολυμέσων ISO όπως MP4. Συγκεκριμένα μέσω του MP4Box υπάρχει η δυνατότητα εισαγωγής και αφαίρεσης ήχου, βίντεο και υποτίτλων σε ένα αρχείο. Τα δεδομένα αυτά μπορεί να βρίσκονται σε διαφορετικές μορφές. Χρησιμοποιείται για αποκωδικοποίηση και αποκωδικοποίηση δεδομένων, καθώς επίσης και για την προσάρτηση μεταδεδομένων σε

υπάρχουσες ροές ή σε ολόκληρο αρχείο μορφής ISO για την παραγωγή αρχείων MPEG. Το MP4Box έχει την ιδιότητα να παρουσιάζει δεδομένα HTTP Adaptive Streaming. Αρχικά τα κωδικοποιημένα βίντεο μετατρέπονται σε .mp4 αρχεία μέσω του ffmpeg και το MP4Box δημιουργεί τα τμήματα του βίντεο και το αντίστοιχο MPD αρχείο ως εξής:

MP4Box -dash num1 -frag num2 -rap -segment-name segment_ output.mp4

- Η παράμετρος -dash num, τμηματοποιεί το αρχείο εισόδου καθορίζοντας την διάρκεια κάθε τμήματος (segment duration) όσο και η τιμή του num1. Το num1 είναι σε ms.
- Η παράμετρος -frag num2 δημιουργεί υποτμήματα εντός των τμημάτων. Ο αριθμός num2 καθορίζει την διάρκεια σε ms, που θα έχουν τα υποτμήματα. Αν το num1 έχει την ίδια τιμή με το num2, τότε υπάρχει μόνο ένα subsegment ανά segment.
- Η παράμετρος -rap έχει στόχο τα τμήματα να ξεκινούν από τυχαία σημεία, όπως τα keyframes. Η διάρκεια του τμήματος μπορεί να διαφέρει ανάλογα με το που βρίσκονται τα keyframes στο βίντεο.
- Μέσω της παραμέτρου -segment-name καθορίζεται το όνομα των τμημάτων, από ένα αυξανόμενο αριθμό, ο οποίος προστίθεται στο όνομα που δίνεται (π.χ segment_1.m4s, segment_2.m4s)
- Η τελευταία παράμετρος output.mp4 καθορίζει το βίντεο που πρόκειται να τμηματοποιηθεί.

Σαν έξοδο έχουμε ένα video representation σε μορφή τμημάτων, ένα αρχικό τμήμα, το οποίο είναι ένα .mp4 αρχείο και ένα MPD. Στην περίπτωση που επιλέξουμε η διάρκεια ενός τμήματος (segment duration) να είναι μικρή, τότε το βίντεο έχει υψηλή αλλαγή ευκρίνειας στο επίπεδο του τμήματος και είναι κατάλληλο για ζωντανή μετάδοση (live), όμως θα υπάρχει μεγάλος αριθμός m4s αρχείων και URL, καθώς και σταθερό μέγεθος request. Αντίθετα, εάν η διάρκεια του τμήματος είναι μεγάλη, τότε ο αριθμός των m4s αρχείων και των URLs θα είναι μικρότερος, τα request sizes θα είναι πιο ευέλικτα και θα βελτιώνεται η απόδοση της cache.

Η διαδικασία και οι εντολές του MP4Box μπορούν να επαναληφθούν με διαφορετικά bitrate, άρα κάθε φορά δημιουργούνται διαφορετικές αναπαραστάσεις (representations). Κάθε αναπαράσταση δημιουργεί ένα διαφορετικό MPD αρχείο. Οι αντίστοιχες εντολές υπάρχουν και για αρχεία ήχου.

Μετά την διαδικασία αυτή, για να ελεγχθεί αν το περιεχόμενο του MPD αρχείου έχει την σωστή μορφή και συνάδει με τις απαιτήσεις του Dash προτύπου, μπορεί να χρησιμοποιηθεί το εργαλείο DASH-IF Validator. Δίνοντας σαν είσοδο το URL του MPD αρχείου (.mpd), στον Validator, μπορεί να γίνει είτε η επαλήθευση του έγκυρου MPD, είτε η παροχή μιας λίστας με προβλήματα εάν υπάρχουν [17].

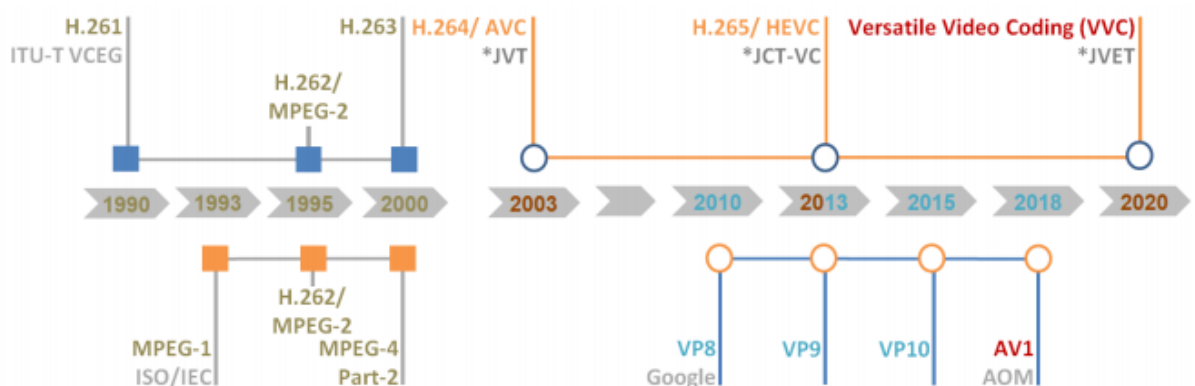
2.6 Κωδικοποιητές

Ο όρος codec προέρχεται από την σύντμηση των όρων coder-decoder, κωδικοποιητής και αποκωδικοποιητής αντίστοιχα, και αναφέρεται σε λογισμικά που χρησιμοποιούν αλγόριθμους συμπίεσης με σκοπό να μειώσουν το μέγεθος του βίντεο και να επιτυγχάνουν streaming μέσω διαδικτύου, αποθήκευση και μετάδοση του. Αυτή η διαδικασία μπορεί να επιτευχθεί σε τηλεοράσεις, ηλεκτρονικούς υπολογιστές, DVD, CD, Blu-ray players, κινητά τηλέφωνα κτλ. Όταν το συμπιεσμένο βίντεο φτάσει την συσκευή του χρήστη, αποσυμπίεζεται και αποκωδικοποιείται ώστε να προβληθεί στην οθόνη. Μέσω της κωδικοποίησης έχουμε απώλεια κάποιων πληροφοριών από τις εικόνες του βίντεο, με αποτέλεσμα να επηρεάζεται η ποιότητα του και τα χαρακτηριστικά του [18].

Οι βίντεο codec εκμεταλλεύονται το χρονικό πλεονασμό (temporal redundancy), μεταξύ των πλαισίων/εικόνων, για παράδειγμα στην περίπτωση που μόνο ένα αντικείμενο αλλάζει σε δύο διαδοχικά frame. Αν ανάμεσα σε δύο διαδοχικά frame, υπάρχουν πολλές περιοχές τις εικόνες που παραμένουν στατικές, τότε μπορεί να αποφευχθεί η μετάδοση πληροφοριών και ο αποκωδικοποιητής θα χρησιμοποιήσει τις τιμές των pixel από τα προηγούμενα frame. Επίσης εκμεταλλεύονται και τον χωρικό πλεονασμό (spatial redundancy), καθώς υπάρχει μεγάλη πιθανότητα γειτονικά pixel να έχουν παρόμοιες ή και ίδιες τιμές. Στην περίπτωση που όλα τα pixel της εικόνας μοιράζονται την ίδια τιμή, έχουμε καλή συμπίεση, αφού στον αποκωδικοποιητή στέλνεται μόνο αυτή η τιμή, αντί οι τιμές όλων των pixel της συγκεκριμένης περιοχής. Τέλος, οι codec λειτουργούν προβλέποντας μικρές περιοχές της εικόνας, που ονομάζονται macroblocks, έτσι μεταδίδουν μόνο τις διαφορές του σήματος ανάμεσα στην προβλεπόμενη και την πραγματική εικόνα.

Στο σχήμα 2.6 παρουσιάζονται τα πρότυπα κωδικοποίησης βίντεο όπως εξελίχθηκαν από το 1990, μέχρι σήμερα. Οι εταιρείες που άνοιξαν τον δρόμο για την ανάπτυξη κωδικοποιητών βίντεο είναι οι ISO/ IEC Moving Picture Experts Group (MPEG) και ITU-T Video Coding Experts Group (VCEG). Οι δύο αυτές εταιρείες συνεργάζονται από

τότε που δημιουργήθηκε το πρότυπο H.264, και αργότερα το H.265, ενώ από το 2017 σχηματίζουν ομάδα με την εταιρεία Video Exploration Team (JVET), με σκοπό την ανάπτυξη του VVC. Παράλληλα η WebM της Google, ανέπτυξε κωδικοποιητές βίντεο VP8 και VP9, την ίδια στιγμή που εμφανίζεται μια πρωτοβουλία της η Alliance for Open Media (AOM) **Error! Reference source not found.**, για την δημιουργία ενός κωδικοποιητή ανοικτού κώδικα, AV1, υποστηρίζοντας ότι έχει την καλύτερη απόδοση. Ως εκ τούτου, απαιτείται η δίκαιη και αντικειμενική σύγκριση των κωδικοποιητών, που παρουσιάζεται πιο κάτω [18]. Στο επόμενο κεφάλαιο, με βάση ένα πείραμα που διεξάγαμε, θα συγκρίνουμε τα πρότυπα κωδικοποίησης x265, VVC, AV1, HM και VP9.



Σχήμα 2.6: Χρονοδιάγραμμα ανάπτυξης προτύπων κωδικοποίησης βίντεο

2.6.1 HEVC

Το πρότυπο κωδικοποίησης high-efficiency video coding (HEVC), είναι ειδικά σχεδιασμένο για κωδικοποίηση υψηλής ανάλυσης HD, και υποστηρίζει μεταφορά βίντεο σε πραγματικό χρόνο (real time transmission). Το HEVC, εγκαινιάστηκε το 2013 από την ομάδα Join Collaborative για την κωδικοποίηση βίντεο, συμπεριλαμβάνοντας μεγάλες βελτιωμένες μεθόδους για την πρόβλεψη κίνησης σε σχέση με τα προηγούμενα πρότυπα κωδικοποίησης τα οποία υποφέρουν από περιορισμένες επιλογές χωρικών αναλύσεων. Για παράδειγμα το H.261, το οποίο υποστηρίζει μορφή ανάλυσης 352x288, που αργότερα επεκτάθηκε σε 1408x1152 σε H.262/MPEG-2. Το H.262/MPEG-2 εξακολουθεί να χρησιμοποιείται σε δορυφορικές τηλεοράσεις και αποθήκευση DVD. Το H.263, εισήχθη το 1995 με στόχο να παρέχει βελτιωμένη ποιότητα βίντεο σε χαμηλότερα

bitrate, επιτρέποντας χαμηλότερη κωδικοποίηση ανάλυσης βίντεο σε sub-QCIF 128x96. Το H.264 ορίζεται από ένα network abstraction layer (NAL) και ένα επίπεδο κωδικοποίησης βίντεο VCL, το οποίο είναι εμπλουτισμένο με νέα εργαλεία κωδικοποίησης μείωσης του bitrate κατά 50% σε σχέση με τις προαναφερθείσες μεθόδους. Το HEVC, χρησιμοποιεί παράλληλη επεξεργασία η οποία συμβάλλει σημαντικά στο HD real time video streaming, έναν δηλαδή από τους πρωταρχικούς στόχους του προτύπου. Επίσης υποστηρίζει ανάλυση βίντεο υψηλής ευκρίνειας Ultra-HD 8Kx4K (7680x4320) και μέχρι 8192x4320 αναλύσεις βίντεο σε σύγκριση με τα 4Kx2K (4096x2048 μέχρι 4096x2304) για aspect ratio 16:9 στο H.264/AVC [20].

2.6.2 x264

Το x264 είναι ένα δωρεάν λογισμικό συμπίεσης βίντεο, η οποία χρησιμοποιείται για την κωδικοποίηση βίντεο σε πρότυπο H.264/AVC – Advanced Video Coding, είναι διαθέσιμο κάτω από τους όρους του GNU General Public License. Το x264 παρέχει μια διεπαφή γραμμής εντολών (command line) και υλοποιεί ένα μεγάλο αριθμό χαρακτηριστικών σε σχέση με άλλους H.264 κωδικοποιητές. Ο κωδικοποιητής είναι αρκετά ισχυρός ώστε να συμπιέζει ψηφιακά βίντεο υψηλής ανάλυσης και να επιτυγχάνει ταυτόχρονα καλή ποιότητα, εξοικονόμηση αποθηκευτικού χώρου, όπως DVD, και μειωμένο ρυθμό μετάδοσης δεδομένων. Παρόλα αυτά η ανάγκη για συμπίεση βίντεο υψηλότερης ανάλυσης καθιστά αυτό τον κωδικοποιητή μια ξεπερασμένη τεχνική [21].

2.6.3 x265

Το x265 είναι ένα λογισμικό συμπίεσης βίντεο βασισμένο στο πρότυπο HEVC- high Efficiency Video Coding. Το HEVC είναι ένας κωδικοποιητής βίντεο, που ανήκει στην ITU-T Video Coding Experts Group, και στόχο έχει να αντικαταστήσει το ευρέως χρησιμοποιημένο AVC (H.264 / MPEG-4). Το HEVC προσφέρει τη διπλάσια συμπίεση των δεδομένων βίντεο, διατηρώντας την ίδια οπτική ποιότητα ή βελτιώνει σημαντικά την ποιότητα του βίντεο στον ίδιο ρυθμό μετάδοσης (bitrate) σε σχέση με το AVC. [22]. Επίσης υποστηρίζει αναλύσεις εικόνas έως 8192x4320, μπορεί να χειριστεί βίντεο μέχρι 8K UHD χρησιμοποιώντας περίπου το μισό αποθηκευτικό χώρο στον δίσκο σε σχέση με το x264. Χρησιμοποιεί ένα δείγμα 8 bit ή 10 bit ανά YUV με συνδυασμό χρωμάτων 4:2:0, 4:2:2 ή 4:4:4. Το x265 υποστηρίζει τα χαρακτηριστικά του x264 και τις λειτουργίες τους, όπως Fixed QP, Fixed Rate Factor, Average bitrate, Video Buffering, και Macroblock-tree rate control. Το Macroblock-tree rate control, ελέγχει την ποιότητα του βίντεο

παρακολουθώντας πόσο συχνά χρησιμοποιούνται τα frames, έτσι ώστε να μπορεί να προβλέψει τα επόμενα frames. Η μέθοδος διευκόλυνε ιδιαίτερα αυτούς που ήδη χρησιμοποιούσαν το H.264, αφού για την αναβάθμιση στο H.265 δεν απαιτούνται πολύ μεγάλες αλλαγές. Το x265 είναι διαθέσιμο είτε μέσω της δεύτερης έκδοσης από το GNU General Public License (GPL), είτε μέσω εμπορικής άδειας όπως και στο x264 [22].

2.6.4 VP9

Η μέθοδος VP9, είναι μια τεχνολογία ανοιχτού κώδικα, που δημιουργήθηκε από την Google [23] με σκοπό να αντικαταστήσει το VP8 [24][25] και διατίθεται δωρεάν, χωρίς κόστος άδειας. Με την δημιουργία του libvpx-VP9, η Google στοχεύει να μειώσει το bandwidth αποθήκευσης και streaming στα βίντεο του YouTube [26]. Ένα από τα χαρακτηριστικά του VP9 είναι ότι είναι προσαρμοσμένο για ανάλυση βίντεο. μεγαλύτερη από 1080p, όπως το UHD, υποστηρίζει βίντεο HDR και επιτρέπει συμπίεση χωρίς απώλειες. Μέσα από έρευνες αποδείχτηκε ότι το VP9, είναι περίπου 30% μέχρι 40% πιο αποδοτικό από το x264. Υπάρχει δυνατότητα πρόσβασης μέχρι και 86,39% σε προγράμματα περιήγησης από κινητά και ηλεκτρονικούς υπολογιστές, σε αντίθεση με το HEVC, που δεν υποστηρίζεται σε προγράμματα περιήγησης. Το VP9 είναι ο μόνος κωδικοποιητής που υποστηρίζει αναπαραγωγή πολυμέσων σε οποιοδήποτε πρόγραμμα, καθώς υποστηρίζει HTML5 video tags, που επιτρέπουν την αναπαραγωγή βίντεο .webm. Είχε παγώσει τον Ιούνιο του 2013, αλλά αργότερα προωθήθηκε από την Google και το Youtube είναι το μόνο που υιοθέτησε το πρότυπο VP9. Από το 2016 η Netflix [27] έδειξε ενδιαφέρον για το VP9, όπου για πρώτη φορά η χρήση του έδειξε εξοικονόμηση εύρους ζώνης κατά μέσο όρο 36%.

2.6.5 VVC

Το VVC (Versatile Video Coding) είναι ένα καινούργιο λογισμικό συμπίεσης βασισμένο στο πρότυπο HEVC, και είναι διάδοχος του H.265. Συχνά αναφέρεται και ως Future Video Coding (FVC) ή ITU-T H.266 [28]. Σύμφωνα με το Γερμανικό ερευνητικό ίδρυμα, το VVC/H.266 μπορεί να μειώσει το ρυθμό μετάδοσης δεδομένων έως και 50% σε σχέση με το HEVC/H.265. Υποστηρίζει αναλύσεις από 4K μέχρι 16K καθώς και βίντεο 360°. Μπορεί να χρησιμοποιηθεί αποτελεσματικά σε κινητά τηλέφωνα για αναπαραγωγή και εγγραφή βίντεο, όπου η χωρητικότητα είναι περιορισμένη, αλλά και σε τηλεοράσεις για αναπαραγωγή βίντεο υψηλής ευκρίνειας. Για παράδειγμα, ένα βίντεο Ultra HD 90 λεπτών που κωδικοποιείται με HEVC/H.265 απαιτεί σήμερα περίπου 10 GB

δεδομένων, ενώ με το VVC/H.266 θα είναι αρκετά περίπου 5 GB **Error! Reference source not found..**

2.6.6 HM

Ο κωδικοποιητής HM, αποτελεί λογισμικό επαλήθευσης που αναφέρεται στο πρότυπο HEVC (test model) , και αναπτύχθηκε από την ομάδα Joint Collaborative Team on Video Coding (JCT-VC) [29]. Είναι κατάλληλο για έλεγχο συμμόρφωσης και πειραματισμό με τα διάφορα χαρακτηριστικά που παρέχει το HEVC, αλλά όχι για έλεγχο υλοποιήσεων πραγματικού χρόνου [30].

2.6.7 AV1

Το AV1, είναι ένας κωδικοποιητής ανοικτού κώδικα και παρέχει δωρεάν άδεια χρήσης. Ξεκίνησε το 2018 από την Alliance of Open Media Video και τον τελευταίο χρόνο έχει βελτιώσει σημαντικά τον χρόνο και την απόδοση της συμπίεσης. Σχεδιάστηκε ως διάδοχος του VP9 της Google και για να ανταγωνιστεί το x265, δημιουργώντας μια μορφή βίντεο που έχει υψηλή απόδοση συμπίεσης και είναι κατάλληλη για χρήση σε προγράμματα περιήγησης και στο διαδίκτυο. Σε δοκιμές του Facebook, οι οποίες προσεγγίζουν τις δοκιμές του πραγματικού κόσμου, ο AV1 πέτυχε κατά 53% μεγαλύτερη συμπίεση δεδομένων, από ότι ο x264. Ένα πλεονέκτημα του AV1, είναι ότι παρέχει δωρεάν άδεια χρήσης σε σχέση με τη HEVC, που απαιτεί πληρωμή των δικαιωμάτων [31]. Είναι κωδικοποιητής ήχου και βίντεο ανοιχτά διαθέσιμος στην αγορά και εύκολα προσβάσιμος για προγραμματιστές υλικού έτσι ώστε να καλύπτει την αυξανόμενη ανάγκη των εφαρμογών βίντεο όπως η τηλεδιάσκεψη και τα ζωντανής μετάδοσης βιντεοπαιχνίδια [21].

Κεφάλαιο 3

Σύγκριση Προτύπων Κωδικοποίησης

3.1	Μεθοδολογία Υλοποίησης	29
3.1.1	Σύνολο Δεδομένων Βίντεο	30
3.1.2	Πειραματική ρύθμιση βίντεο	31
3.2	Αξιολόγηση Αποτελεσμάτων	33
3.2.1	Αλγόριθμος PSNR	34
3.2.2	Ο αλγόριθμος Bjontegaard: BD-Rate Distortion	35
3.3	Αξιολόγηση Κωδικοποιημένων Βίντεο	36
3.4	Συγκεντρωτικά Αποτελέσματα	43
3.5	Σχολιασμός Αποτελεσμάτων	45

Στην παρούσα εργασία, σε πειραματικό στάδιο στόχος ήταν η σύγκριση σύγχρονων κωδικοποιητών εκτός από το x264, έτσι ώστε να δούμε ποιος κωδικοποιητής από τους VVC, AV1, HM, VP9 και x265, είναι ο καλύτερος, υπολογίζοντας την διαφορά που υπάρχει στο Bitrate μεταξύ των κωδικοποιητών για την ίδια ποιότητα βίντεο. Για την σύγκρισή τους, χρησιμοποιήθηκε ο αλγόριθμος αξιολόγησης PSNR. Σε αυτό το κεφάλαιο παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, τα αποτελέσματα και η σύγκρισή τους με αποτελέσματα άλλων μελετών.

3.1 Μεθοδολογία Υλοποίησης

Σε αυτή την ενότητα παρέχονται λεπτομέρειες σχετικά με την προσέγγιση που χρησιμοποιήθηκε για την αντικειμενική σύγκριση της απόδοσης των κωδικοποιητών βίντεο που εξετάστηκαν. Αρχικά παρουσιάζονται τα σύνολα δεδομένων που χρησιμοποιήθηκαν για αυτή την μελέτη και η πειραματική ρύθμιση, που δίνονται από τον πίνακα 3.2 και τον πίνακα 3.3

3.1.1 Σύνολο Δεδομένων Βίντεο

Χρησιμοποιήθηκαν δύο σύνολα δεδομένων στο πλαίσιο της παρούσας μελέτης, ένα σύνολο από έντεκα βίντεο με σκηνές έκτακτης ανάγκης σε περιοχές καταστροφών και ένα γενικό σύνολο βίντεο από το Netflix το οποίο αποτελείται από δέκα ακατέργαστα βίντεο, όπως απεικονίζονται στους πίνακες 3.2,3,3 και στο σχήμα 3.2

Όσο αφορά το σύνολο βίντεο από το Netflix, τα βίντεο έχουν ανάλυση 768x432, διάρκεια 10 δευτερολέπτων και είναι σε μορφή yuv420p raw format. Κάποια στιγμιότυπα από αυτά απεικονίζονται στο σχήμα 3.2a. Τα επτά βίντεο τα οποία χρησιμοποιήθηκαν για το πειραματικό στάδιο έχουν frame rate 25fps, ενώ τα υπόλοιπα τρία βίντεο έχουν frame rate 50fps. Το σύνολο βίντεο από Emergency Scenes αποτελείται από έντεκα βίντεο διαφορετικών σεναρίων όπως απεικονίζονται στο σχήμα 3.2b, τα οποία τραβήχτηκαν κατά την διάρκεια μιας ζωντανής εκδήλωσης του Κυπριακού Συνδέσμου Επείγουσας και Προνοσοκομειακής Αντιμετώπισης. Επτά από αυτά τα βίντεο έχουν ανάλυση 720x576 και τα υπόλοιπα τέσσερα έχουν ανάλυση 1440x1080, όλα έχουν διάρκεια δέκα δευτερολέπτων, frame rate 25fps και βρίσκονται σε ακατέργαστη μορφή yuv420p.

Επιλέχθηκαν βίντεο γενικών χρήσεων από το Netflix, και βίντεο διαφορετικών αναλύσεων από το Emergency Scenes, έτσι ώστε να υπάρχει ποικιλία όσο αφορά την ανάλυση, το framerate και το περιεχόμενο του βίντεο. Με αυτό τον τρόπο επιτυγχάνεται αντικειμενική σύγκριση, καθώς τα πρότυπα συμπίεσης βίντεο είναι ιδιαίτερα ευαίσθητα στις παραπάνω παραμέτρους.

Yuv420p raw format

Η παράμετρος raw αναφέρεται σε μη συμπιεσμένα, ακατέργαστα βίντεο, οπότε δεν υπάρχει διαθέσιμη κεφαλίδα στον κωδικοποιητή για να προσδιορίσει την ανάλυση εισόδου-input resolution, το frame rate, και το colorspace. Το νούμερο 420 αναφέρεται στο colorspace του βίντεο. Η μορφή YUV αναπαριστάτε από 8 bit pixel και αποτελείται από ένα παράγοντα Y, το οποίο αντιπροσωπεύει την φωτεινότητα, χρησιμοποιεί 4 bit και είναι ο πιο αποφασιστικός παράγοντας όσο αφορά την ποιότητα του βίντεο. Οι παράγοντες UV, έχουν 2 bit και αντιστοιχούν στα χρώματα μπλε και κόκκινο αντίστοιχα.

3.1.2 Πειραματική ρύθμιση βίντεο

Η πειραματική ρύθμιση (experimental setup) που χρησιμοποιήθηκε στα πειράματα απεικονίζεται στο πίνακα 3.3. Οι παραμέτροι κωδικοποίησης επιλέχθηκαν με στόχο να επιτρέπουν τη δίκαιη σύγκριση μεταξύ των κωδικοποιητών βίντεο βάσει της σχετικής βιβλιογραφίας[32]-[36].

Σε κάθε βίντεο, πρέπει να οριστούν οι προκαθορισμένες παραμέτροι (preset parameters), που καθορίζουν την ποιότητα. Για κωδικοποίηση x265 ορίστηκε -placebo, δηλαδή πιο αργή κωδικοποίηση, αλλά με την καλύτερη ποιότητα, για κωδικοποίηση VVC, και HM ορίστηκαν τυχαία, Random Access, και για το AV1 και VP9 ορίστηκαν -good και -best αντίστοιχα. Για κάθε πρότυπο κωδικοποίησης χρησιμοποιήθηκαν οι πιο πρόσφατες εκδόσεις λογισμικού, οι οποίες κυκλοφόρησαν τον Φεβρουάριο του 2019. Συγκεκριμένα, χρησιμοποιήθηκε AV1 1.0, HM 16.9, ffmpeg's x265 build και libvpx VP9. Παρόλο που το VP9 δεν είναι ξεπερασμένο, σε αυτή την μελέτη χρησιμοποιείται για σκοπούς συγκριτικής αξιολόγησης καθώς παρουσιάζει μειωμένη αποτελεσματικότητα συμπίεσης σε σχέση με τους υπόλοιπους κωδικοποιητές.

Το κάθε ένα από τα ασυμπίεστα βίντεο, κωδικοποιήθηκε σε VVC, AV1, HM, VP9 και x265, με την βοήθεια του ffmpeg. Για να υπάρχει δίκαιη σύγκριση η παράμετρος κβαντοποίησης αλλάζει ώστε να επιτευχθεί ένα παρόμοιο rate distortion performance για όλα τα πρότυπα. Το κάθε ασυμπίεστο βίντεο κωδικοποιήθηκε σε σταθερής ποιότητας κωδικοποίησης QP set (22,27,32,37) για τα VVC, HM, και x265, ενώ για το AV1 και VP9 κωδικοποιήθηκε σε QP set (27,35,46,55). Με αυτό τον τρόπο λαμβάνεται υπόψη ένα ευρύ φάσμα αντιπροσωπευτικού εύρου ζώνης και υπάρχουν είκοσι κωδικοποιήσεις για το κάθε ασυμπίεστο βίντεο, αφού στο σύνολο υπάρχουν 440 κωδικοποιήσεις.

Η παράμετρος κβαντοποίησης (QP) ελέγχει το μέγεθος της συμπίεσης για κάθε Macroblock σε ένα πλαίσιο. Όταν το QP έχει μεγάλη τιμή, τότε θα υπάρχει μεγάλη κβαντοποίηση, άρα μεγαλύτερη συμπίεση όμως θα έχουμε περισσότερη απώλεια πληροφορίας με αποτέλεσμα την χαμηλότερη ποιότητα του βίντεο. Οι πιο μικρές τιμές σημαίνουν το αντίθετο. Ο καθορισμός ενός σταθερού QP, σημαίνει ότι το bitrate που προκύπτει διαφέρει κατά πολύ ανάλογα με το περιεχόμενο κάθε σκηνής του βίντεο και κατά συνέπεια τη μη αποτελεσματική αποκωδικοποίηση του βίντεο.

Όσο αφορά την υποστήριξη τυχαίας πρόσβασης (Random Access) , εισάχθηκε ένα intra update κάθε 32 πλαίσια για τα βίντεο με 25fps και ένα intra update κάθε 48 πλαίσια για τα βίντεο με 50fps. Το επίπεδο GOP καθορίζεται από την πιο πάνω παράμετρο.

Dataset Name	No. Videos	Video Resolution	Frame Rate	Duration
Netflix Dataset I	7	768x432	25	10s
Netflix Dataset II	3	768x432	50	10s
Emergency Scenery Video Dataset I	7	720x576	25	10s
Emergency Scenery Video Dataset II	4	1440x1080	25	10s

Πίνακας 3.2: Χαρακτηριστικά Βίντεο Κωδικοποίησης

Video Codec	QP	Random Access	Preset Profile
VVC	22,27,32,37	32/48 ^a	Random access
HM			
x265	placebo		
AV1	good		
VP9	best		

Πίνακας 3.3: Πειραματικές Ρυθμίσεις Κωδικοποίησης



(a)



(b)



(c)

Σχήμα 3.2a: Emergency Scenery Videos. Απεικονίζονται διαφορετικές εκπαιδευτικές σκηνές που καταγράφηκαν κατά την διάρκεια Κυπριακού Συνδέσμου Επείγουσας και Προνοσοκομειακής Αντιμετώπισης. Η εικόνα (a) συλλαμβάνει

πυροσβεστικά οχήματα που πλησιάζουν σε σεισμό, η εικόνα (b) απεικονίζει τους πρώτους ανταποκριτές που αναζωογονούν έναν ασθενή σε μια κινητή κλινική που ιδρύθηκε κοντά στο χώρο έκτακτης ανάγκης και στην εικόνα (c) παρουσιάζεται η εκκένωση των τραυματιών που λαμβάνει χώρα μετά από επείγοντα περιστατικά πυρκαγιάς ή σεισμού.



(α)



(β)



(γ)



(δ)



(ε)

Σχήμα 3.2b: Netflix video dataset. Απεικονίζονται διαφορετικές εικόνες, οι οποίες παρουσιάζουν διαφορετικά γεγονότα. Στις εικόνες (α)-(δ) παρουσιάζονται διαφορετικό γεγονός και αντιπροσωπεύει διαφορετικό περιεχόμενο βίντεο. Στα βίντεο που αφορούν τα (δ) και (ε) υπάρχει γρήγορη κίνηση και μετά εστίασης της κάμερας σε ένα συγκεκριμένο σημείο, ενώ στα υπόλοιπα βίντεο υπάρχει πιο αργή κίνηση.

3.2 Αξιολόγηση Αποτελεσμάτων

Για την αξιολόγηση της ποιότητας του βίντεο, χρησιμοποιήθηκε ο αλγόριθμος αντικειμενικής αξιολόγησης PSNR. Ο αλγόριθμος εφαρμόζεται μεταξύ των πλαισίων του βίντεο πριν την κωδικοποίηση και των αντίστοιχων πλαισίων του βίντεο μετά την κωδικοποίηση του. Επίσης για την σύγκριση του PSNR με το Bitrate χρησιμοποιήθηκε ο αλγόριθμος Bjontegaard: BD-Rate Distortion.

3.2.1 Αλγόριθμος PSNR

Ο όρος PSNR προέρχεται από την Αγγλική Ορολογία Peak Signal-to-Noise Ratio και είναι ο λόγος της μέγιστης δυνατής ισχύος του σήματος ως προς το θόρυβο. Αρχικά είχε αναπτυχθεί για αξιολόγηση ποιότητας εικόνας, όμως τώρα, θεωρείται ο πιο διαδεδομένος και γρήγορος από υπολογιστικής άποψης αλγόριθμος αξιολόγησης βίντεο και έχει μονάδα μέτρησης του είναι τα decibel (dB). Στην περίπτωση αυτή το σήμα είναι το αρχικό βίντεο και ο θόρυβος είναι το σφάλμα που εισάγεται εξαιτίας της συμπίεσης, οπότε το PSNR, εκφράζει την ποιότητα του νέου βίντεο μετά την απωλεστική συμπίεση.

Αρχικά πρέπει να υπολογιστεί το μέσο τετραγωνικό σφάλμα (mean-squared error - MSE) μεταξύ του αρχικού και του συμπιεσμένου/μεταδιδόμενου βίντεο σε επίπεδο πλαισίων. Δεδομένου ενός ασυμπίεστου βίντεο αναφοράς I , και ένα αλλοιωμένο βίντεο K το οποίο προβλήθηκε στον παραλήπτη, με χωρικές διαστάσεις $m \times n$, το MSE υπολογίζεται ως εξής:

$$MSE = \frac{1}{nm} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \quad (1)$$

Οι μεταβλητές i, j εκφράζουν το εικονοστοιχείο (pixel) στη θέση της $m \times n$ εικόνας. Ο MSE παίρνει την τιμή μηδέν, αν και μόνο αν οι δύο εικόνες είναι πανομοιότυπες, και ισχύει ότι $I(i,j) = K(i,j)$.

Το τελικό PSNR είναι αντιστρόφως ανάλογο του MSE και υπολογίζεται ως εξής:

$$PSNR = 10 \log_{10} \frac{MAX_i^2}{MSE} \quad (2)$$

Όπου το MAX_i ορίζεται ως η μέγιστη τιμή που μπορεί να πάρει ένα pixel, δηλαδή το εύρος των επιτρεπόμενων φωτεινότητων και ισούται με $2^n - 1$ όπου το n αντιστοιχεί στον αριθμό των bits ανά εικονοστοιχείο. Για παράδειγμα αν αναφερόμαστε σε grayscale εικόνες αναπαρίστανται με 8bits ανά pixel και η τιμή του MAX_i ορίζεται στο 255. Όσο μεγαλύτερο είναι το PSNR, τόσο καλύτερη είναι η ποιότητα του βίντεο, και αν έχουμε PSNR μεγαλύτερο από 30 dB η αλλοίωση του βίντεο δεν γίνεται αντιληπτή από το ανθρώπινο μάτι.

Ο εναλλακτικός αλγόριθμος PSNR

Στην περίπτωση της μελέτης ο μέσος όρος PSNR, υπολογίζεται σε κάθε πλαίσιο και στη συνέχεια υπολογίζεται κατά μέσο όρο σε ολόκληρο το βίντεο. Συγκεκριμένα υιοθετήθηκε μια δημοφιλής εναλλακτική λύση στον κλασσικό υπολογισμό που έχει περιγραφεί πιο πάνω. Ουσιαστικά, εισήχθησαν τα κατάλληλα βάρη για την φωτεινότητα PSNR_y και το χρώμα PSNR_u και PSNR_v, έτσι ώστε να αντανακλούν καλύτερα την ποιότητα της εικόνας [18]. Η ποιότητα του βίντεο αξιολογήθηκε ως εξής:

$$PSNR_{611} = (6 \times PSNR_y + PSNR_u + PSNR_v) / 8 \quad (3)$$

όπου το Y αντιστοιχεί στην φωτεινότητα, και το U,V αντιστοιχούν στο μπλε και το κόκκινο αντίστοιχα. Οπότε,

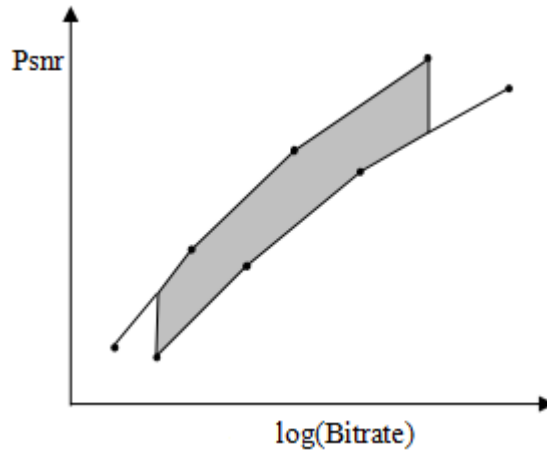
$$PSNR_{Y,U,V} = 10 \log_{10} \frac{MAX_i^2}{MSE} \quad (4)$$

Χρησιμοποιήθηκε το εναλλακτικό PSNR₆₁₁ (global PSNR) γιατί είναι καλύτερο από το κλασσικό PSNR, όσο αφορά την συμπίεση βίντεο, λόγω του ότι η φωτεινότητα αποτελεί τον κύριο παράγοντα στην αξιολόγηση βίντεο. Αφού γίνεται αναφορά σε βίντεο yuv420, είναι πιο αντικειμενικό να χρησιμοποιηθεί η βαρύτητα 611, αντίστοιχα στις τιμές yuv, παρά το κλασσικό PSNR.

3.2.2 Ο αλγόριθμος Bjontegaard: BD-Rate Distortion

Από το 2008 χρησιμοποιείται ευρέως ο αλγόριθμος Bjontegaard, μέσω του οποίου υπολογίζεται η μέση διαφορά bitrate για ισοδύναμη αντικειμενική ποιότητα (τυπικά το PSNR), σε σύγκριση με την απόδοση δύο κωδικοποιητών. Ο αλγόριθμος αυτός συνήθως εμφανίζεται με την συντομογραφία BD-Rate και η βασική ιδέα είναι να χρησιμοποιείται το PSNR Vs log(Bitrate) σε μια γραφική παράσταση, rate distortion graph. Το PSNR είναι ήδη σε λογαριθμική μορφή, οπότε αν λογαριθμιστεί και το bitrate, οι γραμμές της γραφικής παράστασης θα μοιάζουν με ευθείες και θα υπάρχει μια πιο αντικειμενική αξιολόγηση. Οι συντελεστές BD επιτρέπουν την δυνατότητα σύγκρισης οποιωνδήποτε δύο πρότυπων κωδικοποίησης βίντεο. Ένα παράδειγμα παρουσιάζεται στο σχήμα 3.2.2 όπου Η σκιασμένη περιοχή μεταξύ των σημείων αντιπροσωπεύει την διαφορά μεταξύ

των δύο καμπυλών **Error! Reference source not found.** Κάθε BD-rate καμπύλη αντιστοιχεί σε πολυώνυμο τρίτης τάξης και υπολογίζεται η περιοχή ανάμεσα στις δύο καμπύλες. Η περιοχή αυτή αντιστοιχεί στην μέση διαφορά bitrate, για ισοδύναμη ποιότητα PSNR ανάμεσα στα δύο πρότυπα κωδικοποίησης που συγκρίνονται **Error! Reference source not found.**



Σχήμα 3.2.2: PSNR Vs Log(Bitrate), Rate Distortion Graph**Error! Reference source not found.**

Για να υπολογιστεί η εξοικονόμηση bitrate υπολογίζετε το ποσοστό εξοικονόμησης για την αντίστοιχη ποιότητα βίντεο. Αυτό επιτυγχάνεται χρησιμοποιώντας τον αλγόριθμο BD-rate, ο οποίος υπολογίζει αντικειμενικές διαφορές μεταξύ δυο καμπυλών παραμόρφωσης ρυθμού (rate-distortion curves) και δίνει τη ποσοστιαία διαφορά του bitrate. Οι καμπύλες ρυθμού παραμόρφωσης κατασκευάστηκαν βασισμένες στο Y-PSNR και το τελικό ποσοστό διαφοράς υπολογίζεται από τον μέσο όρο στο σύνολο των δεδομένων που εξετάζονται.

3.3 Αξιολόγηση Κωδικοποιημένων Βίντεο

Δημιουργήθηκε ένας rate-distortion πίνακας τιμών για κάθε μια από τις κωδικοποιήσεις με τα πιο πάνω πρότυπα. Ουσιαστικά χρησιμοποιήσαμε το μέσο (Median) και το μέσο όρο (Average) για όλα τα βίντεο για το κάθε QP ξεχωριστά. Το rate αντιστοιχεί στο απαιτούμενο εύρος, bitrate και το distortion αντιστοιχεί στο PSNR. Το PSNR αποτυπώνει την αλλοίωση της ποιότητας του βίντεο σε σχέση με το αρχικό βίντεο πριν την συμπίεση. Οι πίνακες τιμών όσο αφορά το bitrate, το λογαριθμισμένο bitrate και το PSNR για κάθε

κωδικοποιητή και το αντίστοιχο QP set εμφανίζονται στους πίνακες 3.3.1, 3.3.2, 3.3.3, 3.3.4 πιο κάτω.

Rate-Distortion									
Codec	VVC			HM			x265		
QP	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)
22	985.954	2.994	43.914	1180.009	3.072	42.631	1698.900	3.072	43.165
27	522.762	2.718	41.191	600.725	2.779	39.912	906.520	2.779	40.577
32	279.466	2.446	38.430	310.872	2.493	37.317	498.210	2.493	37.951
37	148.348	2.171	35.965	166.954	2.223	34.579	271.320	2.223	35.124

Rate-Distortion						
Codec	AV1			VP9		
QP	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)
27	1306.852	3.116	43.869	1290.166	3.111	42.715
35	829.927	2.919	42.054	810.865	2.909	40.835
46	429.727	2.633	39.128	440.323	2.644	38.101
55	255.630	2.408	36.800	298.865	2.475	36.422

Πίνακας 3.3.1: Netflix video DataSet I: 768x432@25Hz

Πίνακας 3.3.2: Netflix video DataSet I: 768x432@25Hz

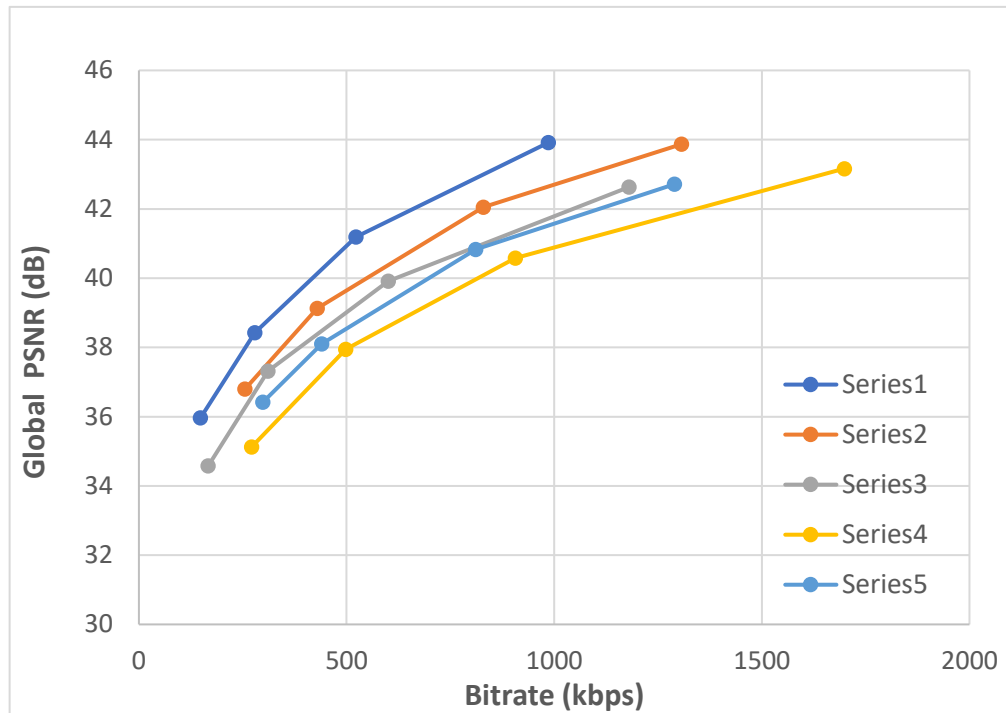
Rate-Distortion									
Codec	VVC			HM			x265		
QP	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)
22	1777.26	3.250	40.38	2434.341	3.386	39.82	3955.780	3.597	40.55
27	786.194	2.896	38.22	1028.537	3.012	37.21	1463.370	3.165	37.83
32	427.035	2.630	35.64	520.454	2.716	34.48	678.750	2.832	35.03

Rate-Distortion						
Codec	AV1			VP9		
QP	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)	Bitrate (Kbps)	Bitrate (Log ₁₀)	PSNR (db)
27	3027.241	3.481	40.98	3073	3.488	40.23
35	1716.185	3.235	39.73	1844	3.266	38.84
46	802.601	2.904	37.41	1012	3.005	36.74
55	510.987	2.708	35.23	665	2.823	34.96

Πίνακας 3.3.3: Netflix video DataSet II: 768x432@50Hz

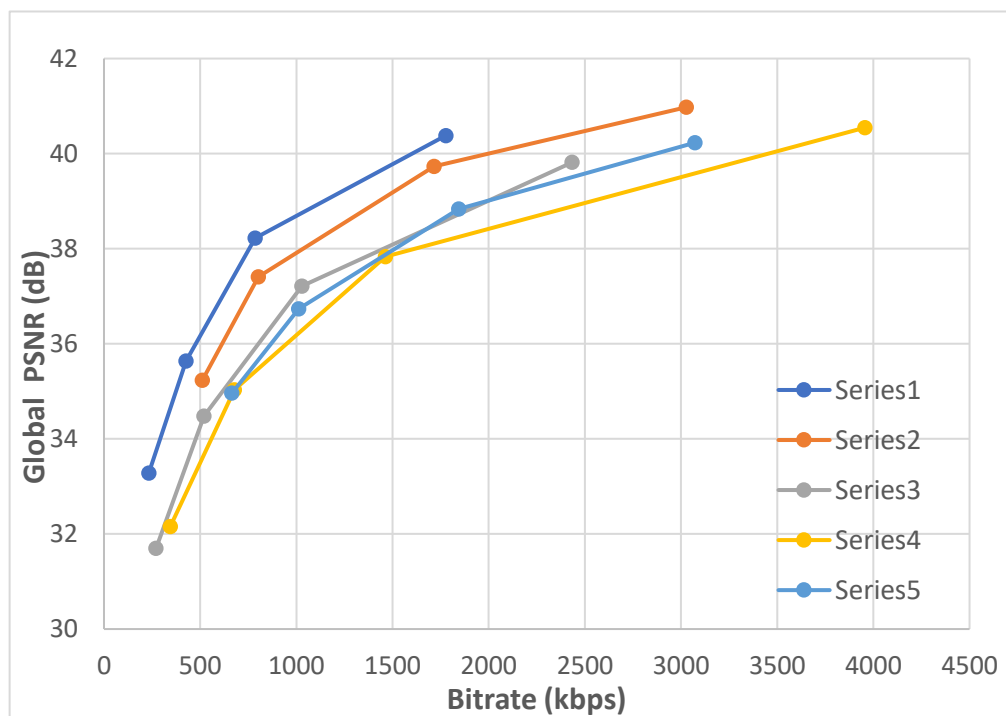
Πίνακας 3.3.4: Netflix video DataSet II: 768x432@50Hz

Στο σχήμα 3.3.5 παρουσιάζεται η γραφική που προκύπτει από τα κωδικοποιημένα βίντεο



με frame rate 25, και στο σχήμα 3.3.6 παρουσιάζεται η γραφική που προκύπτει από τα κωδικοποιημένα βίντεο με frame rate 50.

Σχήμα 3.3.5: Netflix Video Dataset I: PSNR Vs Bitrate, @25Hz



Σχήμα 3.3.6: Netflix Video Dataset II: PSNR Vs Bitrate, @50Hz

Για την κωδικοποίηση x265 μπορούμε να πάρουμε απευθείας το global PSNR και το bitrate από τον ffmpeg, όπως φαίνεται στο σχήμα 3.3.7.

```
encoded 217 frames in 405.15s (0.54 fps), 3428.68 kb/s, Avg QP:23.18,  
Global PSNR: 42.447, SSIM Mean Y: 0.9901920 (20.084 dB)
```

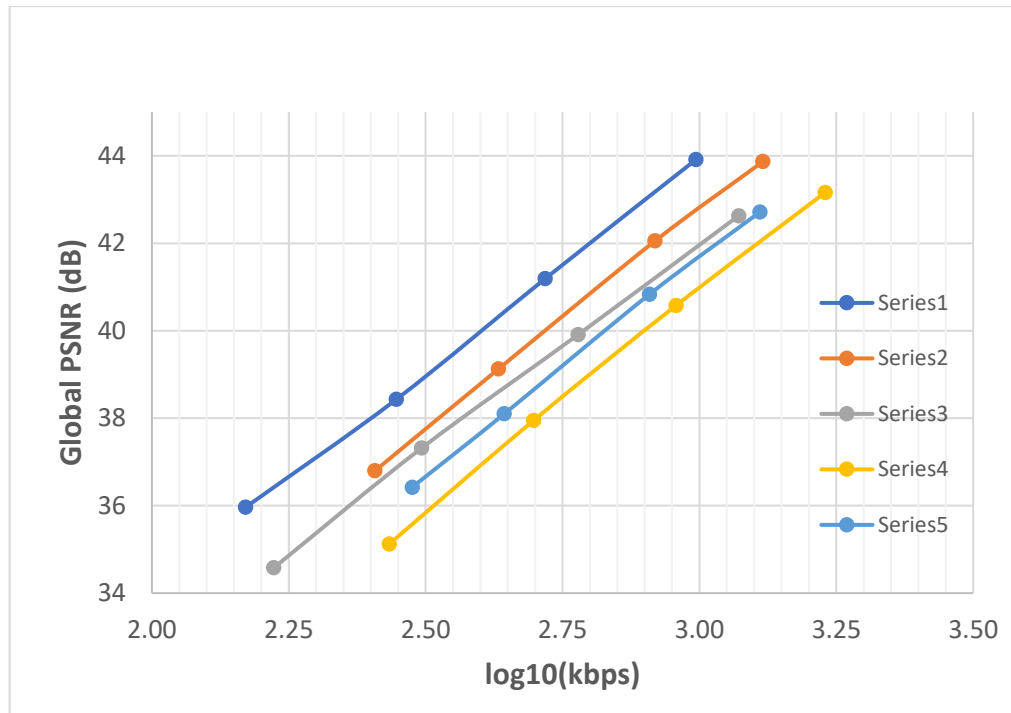
Σχήμα 3.3.7: x265, Bitrate & Global PSNR

Για τα υπόλοιπα πρότυπα κωδικοποίησης παίρνουμε το bitrate σε kbps από τον ffmpeg. Όσο αφορά το PSNR, κατά την κωδικοποίηση υπολογίζεται το PSNR για το κάθε ένα συστατικό Y, U και V ξεχωριστά, όπως φαίνεται στο σχήμα 5.8. Επεξεργασόμαστε αυτά τα δεδομένα με την φόρμουλα $(6*Y+U+V)/8$ για να πάρουμε το νέο global PSNR.

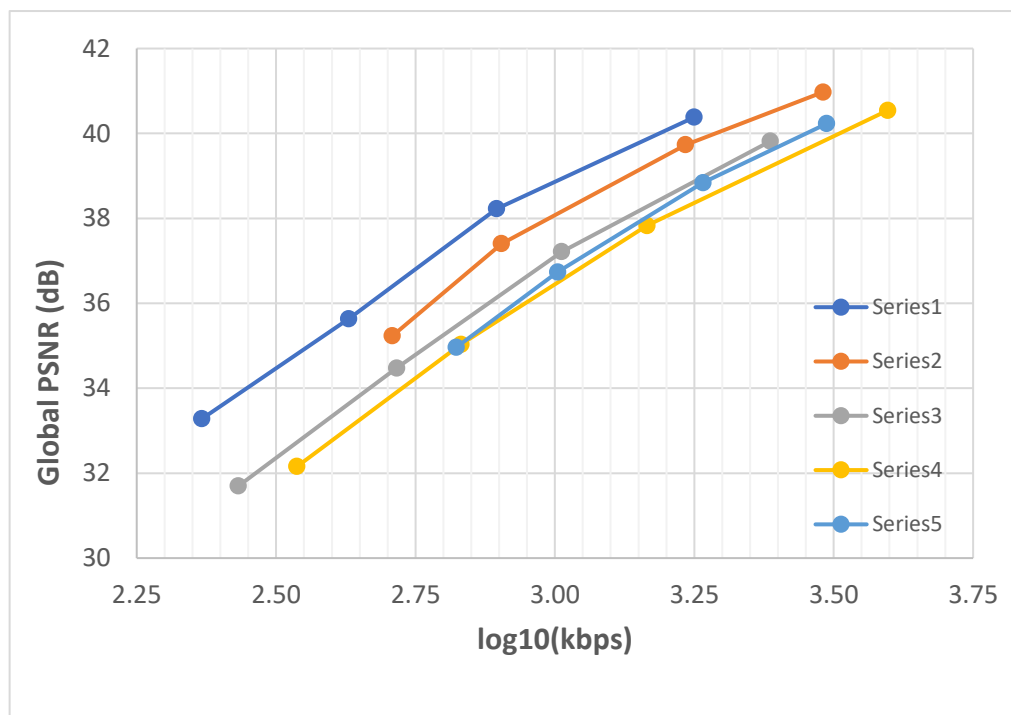
Total Frames		Bitrate	Y-PSNR	U-PSNR	V-PSNR	YUV-PSNR
217	a	1579.5152	41.7374	39.0995	40.9174	40.6746

Σχήμα 3.3.8: VVC, Bitrate & Y-PSNR, U-PSNR, V-PSNR

Μετά από την φόρμουλα, το PSNR, είναι ήδη λογαριθμισμένο, και στη συνέχεια λογαριθμίζεται και το bitrate για καλύτερη επεξεργασία των αποτελεσμάτων. Στο σχήμα 3.3.9 παρουσιάζεται η γραφική που προκύπτει από τα κωδικοποιημένα βίντεο με frame rate 25 με λογαριθμισμένο bitrate και psnr, και στο σχήμα 3.3.10 παρουσιάζεται η γραφική που προκύπτει από τα κωδικοποιημένα βίντεο με frame rate 50, με λογαριθμισμένο bitrate και psnr. Οι πιο κάτω γραφικές παραστάσεις δημιουργήθηκαν με βάση το μέσο (Median), για κάθε κωδικοποίηση και QP ξεχωριστά.



Σχήμα 3.3.9 Netflix Video Dataset I: Rate Distortion Curves (PSNR Vs $\log_{10}(\text{Bitrate})$ of Mean Values) @25Hz



Σχήμα 3.3.10: Netflix Video Dataset II: Rate Distortion Curves (PSNR Vs $\log_{10}(\text{Bitrate})$ of Mean Values) @50Hz

Ο πίνακας τιμών rate-distortion μεταφέρθηκε στο Excel αρχείο BitRate-Gains. Το συγκεκριμένο αρχείο χρησιμοποιώντας τον αλγόριθμο BD-Rate ,υπολογίζει την εξοικονόμηση σε εύρος ζώνης για ισοδύναμη ποιότητα εικόνας μεταξύ των διαφόρων κωδικοποιήσεων. Τα αποτελέσματα εμφανίζονται υπό ποσοστιαία μορφή κάτω από το %BITtab, όπως παρουσιάζονται στους πίνακες 3.3.11 και 3.3.12.

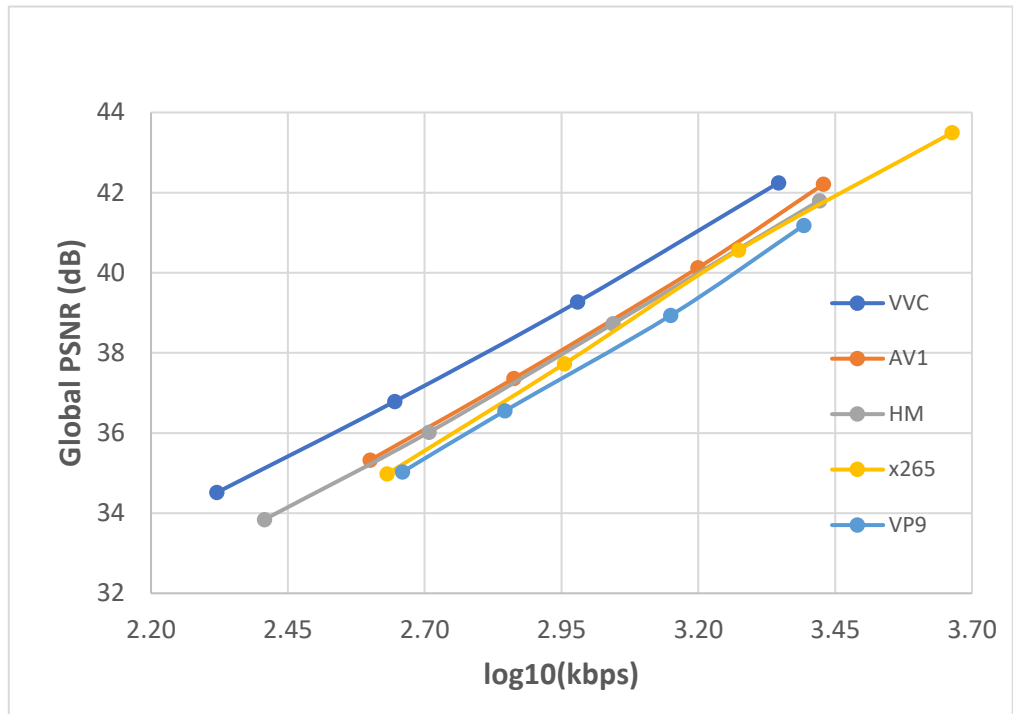
%BitRate Distortion				
Codecs	AV1	HM	x265	VP9
VVC	-26.948	-34.836	-40.240	-46.499
AV1		-10.837	-18.654	-28.248
HM			-8.920	-19.507
x265				-11.921

Πίνακας 3.3.11: %BitRate Distortion, Netflix Video DataSet I: 768x432@25Hz

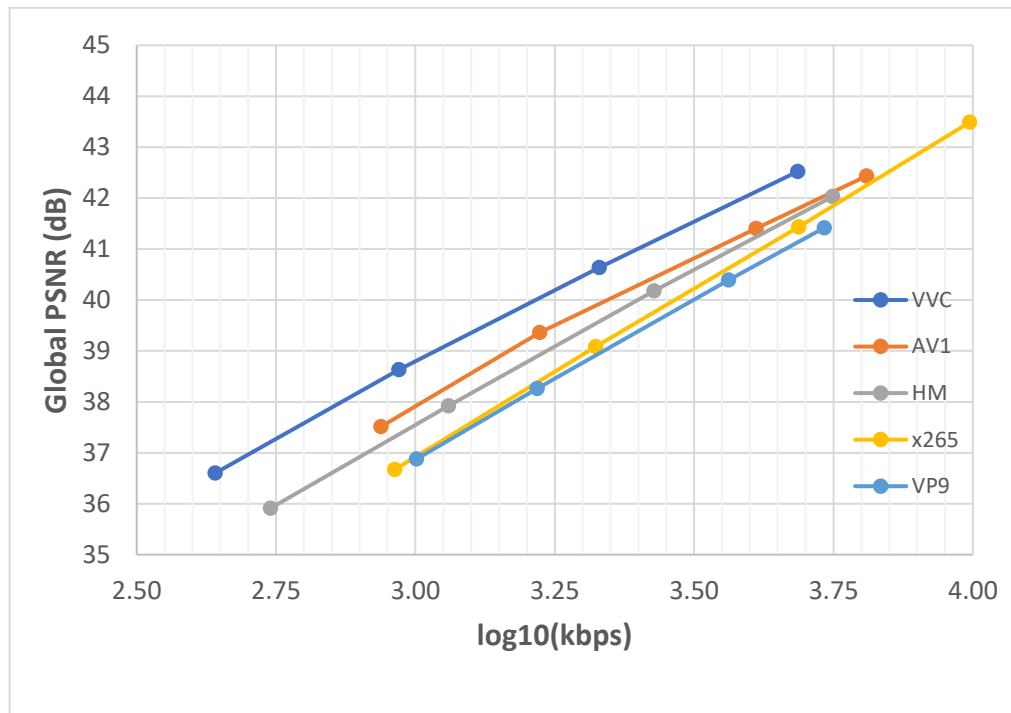
%BitRate Distortion				
Codecs	AV1	HM	x265	VP9
VVC	-22.223	-34.266	-44.955	-43.084
AV1		-16.864	-30.637	-27.110
HM			-15.377	-12.352
x265				4.732

Πίνακας 3.3.12: %BitRate Distortion, Netflix Video DataSet II: 768x432@50Hz

Τα συγκεντρωμένα αποτελέσματα των επτά βίντεο που χωρίζουν το πρώτο σύνολο δεδομένων βίντεο σκηνικού έκτακτης ανάγκης, με ανάλυση βίντεο 720×576 , παρουσιάζονται στην γραφική παράσταση 3.3.13 και στον πίνακα 3.3.15. Η γραφική παράσταση 3.3.14 και ο πίνακας τιμών 3.3.16 απεικονίζουν τα αποτελέσματα του δεύτερου συνόλου δεδομένων που αποτελείται από τέσσερα βίντεο με ανάλυση βίντεο 1440×1080 .



Πίνακας 3.3.13: Emergency Scenery Video Dataset I: Rate Distortion Curves (PSNR Vs $\log_{10}(\text{Bitrate})$ of Mean Values), Resolution: 720x576



Πίνακας 3.3.14: Emergency Scenery Video Dataset II: Rate Distortion Curves (PSNR Vs $\log_{10}(\text{Bitrate})$ of Mean Values), Resolution: 1140x1080

%BitRate Distortion				
Codecs	AV1	HM	x265	VP9
VVC	-26.880	-31.913	-37.937	-44.784
AV1		-5.700	-14.807	-22.663
HM			-10.053	-18.570
x265				-7.845

**Πίνακας 3.3.15: %BitRate Distortion, Emergency Scenery Video Dataset I:
720x576@25Hz**

%BitRate Distortion				
Codecs	AV1	HM	x265	VP9
VVC	-25.784	-37.407	-44.408	-49.862
AV1		-8.560	-23.927	-31.306
HM			-10.699	-19.531
x265				-7.560

**Πίνακας 3.3.16: %BitRate Distortion, Emergency Scenery Video Dataset II:
1140x1080@25Hz**

3.4 Συγκεντρωτικά Αποτελέσματα

Από τους πίνακες τιμών 3.3.1, 3.3.2, 3.3.3, 3.3.4, που αφορούν το σύνολο δεδομένων βίντεο του Netflix παρατηρούμε ότι όταν το QP είναι ίσο με 22, έχουμε πιο μεγάλο PSNR και Bitrate από ότι τα υπόλοιπα QP. Παρατηρούμε ότι όσο αυξάνεται το QP, τόσο μικρότερο PSNR έχουμε, γιατί λόγω του ότι έχουμε μεγάλη κβαντοποίηση, έχουμε μεγαλύτερη απώλεια πληροφορίας. Αυτό έχει ως αποτέλεσμα να χάνονται σημαντικές λεπτομέρειες του βίντεο, και να έχουμε αλλοιωμένη ποιότητα του βίντεο, έτσι το αρχικό και το τελικό βίντεο διαφέρουν και να έχουμε μικρό PSNR.

Από τα σχήματα 3.3.5 και 3.3.6 οι γραφικές PSNR Vs Bitrate είναι καμπύλες, με το x265 να έχει μεγαλύτερο bitrate από τα υπόλοιπα. Παρατηρούμε ότι το VP9 και το AV1 ξεκινούν με μεγαλύτερο bitrate από ότι τα υπόλοιπα. Από τα σχήματα 3.3.9 και 3.3.10 παρατηρούμε ότι οι γραφικές με το λογαριθμισμένο bitrate είναι ευθείες. Στο σχήμα 3.3.9 για Framerate=25 καθώς αυξάνεται η ανάλυση βίντεο, το AV1 τείνει να ξεπερνά τόσο

το HM όσο και το x265. Στο σχήμα 3.3.6 για Framerate=50, το x265 ξεπερνά κατά πολύ το VP9 της Google. Φαίνεται ότι όσο πιο χαμηλό PSNR έχουμε, τόσο πιο χαμηλό είναι το Bitrate, και το αντίθετο. Όσο αυξάνεται το Bitrate, τόσο πιο καλή είναι η ποιότητα του βίντεο.

Από τους πίνακες του BD-Rate, των πινάκων 3.3.11 και 3.3.12, μπορούμε να εξάγουμε συμπεράσματα όσο αφορά την διαφορά που υπάρχει στο Bitrate μεταξύ των κωδικοποιητών για την ίδια ποιότητα βίντεο. Για παράδειγμα όταν το framerate είναι 50 το %BIT μεταξύ VVC Vs AV1 είναι -22.223. Αυτό σημαίνει ότι το VVC για την ίδια ποιότητα βίντεο χρησιμοποιεί 22,223 πιο λίγο bitrate από ότι το AV1. Από τους πίνακες βλέπουμε ότι ο VVC είναι ο καλύτερος από τους κωδικοποιητές που συγκρίνουμε, και ακολουθείται από τους AV1 και HM. Το VP9 θεωρείται ήδη ξεπερασμένο, ενώ το AV1 επιτυγχάνει καλύτερη απόδοση από τις δύο τυποποιήσεις του HEVC, δηλαδή το HM και το x265, με τον AV1 να μειώνει τις απαιτήσεις του bitrate κατά 16.8% και 30.6 % αντίστοιχα.

Όσο αφορά τα βίντεο έκτακτης ανάγκης, τα αποτελέσματα τεκμηριώθηκαν και για τα δύο σύνολα δεδομένων, ανάλυσης 720x576 και 1140x1080. Τα δύο σύνολα δεδομένων μοιράζονται τα ίδια χαρακτηριστικά βίντεο εκτός από την ανάλυση βίντεο, έτσι κάποιος μπορεί να συμπεράνει ότι το VVC και το AV1 επιτυγχάνουν ένα από τους πρωταρχικούς τους στόχους, την επίτευξη μεγαλύτερης αποτελεσματικότητας συμπίεσης για ψηλές και εξαιρετικά ψηλές (ultra-high) αναλύσεις βίντεο.

Όπως φάνηκε και από τα αποτελέσματα από το σύνολο δεδομένων του Netflix, η απόδοση κωδικοποίησης του VVC, υπερέχει σε σχέση με τα άλλα πρότυπα συμπίεσης, και για τα δύο σύνολα δεδομένων. Τόσο το VVC, όσο και το AV1 αυξάνουν το κέρδος Bitrate, σε σύγκριση με τους κωδικοποιητές που εξετάστηκαν για το σύνολο δεδομένων υψηλότερης ανάλυσης. Από την ερμηνεία των αποτελεσμάτων, αντιλαμβανόμαστε ότι όσο αφορά την απόδοση κωδικοποίησης το AV1 πλέον ξεπερνά σταθερά τους κωδικοποιητές HEVC, δηλαδή τόσο το HM, όσο και το x265, και μειώνει τις απαιτήσεις bitrate, έναντι του VVC. Από την άλλη πλευρά, ο HM μειώνει τις απαιτήσεις bitrate, σε σχέση με το x265 και το VP9 περίπου το ίδιο και για τα δύο σύνολα δεδομένων, κατά 10% και 19% αντίστοιχα. Το ίδιο ισχύει για το x265 σε σύγκριση με το VP9, όπου το bitrate μειώνεται περίπου κατά 7.5%. Με άλλα λόγια, δεν υπάρχει ανάλυση βίντεο, όπου το bitrate να αυξάνει τα κέρδη του, στους δύο κωδικοποιητές που βασίζονται στο HEVC.

3.5 Σχολιασμός αποτελεσμάτων

Τα αποτελέσματα από το σύνολο δεδομένων του Netflix I, είναι συγκρίσιμα με τα αποτελέσματα του συνόλου δεδομένων βίντεο σκηνικών έκτακτης ανάγκης I, καθώς και τα δύο σύνολα δεδομένων έχουν παρόμοιες αναλύσεις βίντεο 768x432 (Netflix) έναντι 720x576 (Emergency scenes) και τον ίδιο framerate, 25fps. Παρά το ότι το σύνολο δεδομένων αποτελείτο από ένα περιορισμένο αριθμό βίντεο με framerate 50fps (συγκεκριμένα τρία βίντεο), ένα ενδιαφέρον αποτέλεσμα είναι ότι οι κωδικοποιητές AV1 και VP9 επιτυγχάνουν καλύτερα αποτελέσματα σε σύγκριση με τα αποτελέσματα των βίντεο με 25fps. Συγκεκριμένα, ο AV1 αυξάνει τα κέρδη bitrate σε σύγκριση με τους κωδικοποιητές HEVC και μειώνει τις απαιτήσεις του bitrate έναντι του VVC. Ομοίως, καταγράφηκαν μειωμένα κέρδη bitrate στον κωδικοποιητή VP9, ο οποίος, για πρώτη φορά, δεν είχε κατώτερη απόδοση κωδικοποίησης σε σύγκριση με τον κωδικοποιητή x265.

Σύμφωνα με τα σημερινά δεδομένα, τα παλαιότερα πρότυπα κωδικοποίησης βίντεο αναπτύχθηκαν και λειτουργούν σε μικρότερες χωρικές αναλύσεις. Τα πιο πρόσφατα και αναδυόμενα πρότυπα επικεντρώνονται σε υψηλές αναλύσεις HD (όπως το HEVC), και αναδυόμενες εφαρμογές, όπως η κωδικοποίηση βίντεο 360° (όπως το AV1 και VVC).

Τα πρότυπα x265 και VP9, βελτιστοποιούν το επίπεδο του λογισμικού και ως εκ τούτου πληρούν τις προϋποθέσεις για απόδοση σε πραγματικό χρόνο. Εκτός από αυτά, τα πρότυπα VVC, AV1 και HM απαιτούν μεγάλο χρόνο συμπίεσης δεδομένου ότι η κύρια χρήση τους προορίζεται στο να επικυρωθεί η αποτελεσματικότητα της συμπίεσης που παρέχουν, αντί να χρησιμοποιούνται σε εφαρμογές πραγματικού χρόνου.

Βασισμένοι στην σύγκριση πολλών προτύπων κωδικοποίησης βίντεο διαπιστώνουμε ότι το πρότυπο VVC , δίνει την καλύτερη απόδοση κωδικοποίησης μέχρι σήμερα. Η σύγκριση δείχνει ότι το VVC ξεπερνά το AV1, όπως αναφέρεται στο [18].

Σύγκριση Αποτελεσμάτων

Στο άρθρο IEEE spectrum , πέραν από την πιο πάνω σύγκριση από το σύνολο δεδομένων με βίντεο του Netflix και των σκηνικών έκτακτης ανάγκης, έγινε και σύγκριση με διάφορες κωδικοποιήσεις ιατρικών βίντεο. Συμφωνεί με τα πιο αποτελέσματα που αναφέρονται στην παράγραφο 6.2, ότι δηλαδή το VVC ξεπερνά όλους τους υπόλοιπους κωδικοποιητές, και προσφέρει καλύτερη απόδοση συμπίεσης από το HEVC.

Από την βιβλιογραφία μελετήθηκαν πολλές έρευνες που σύγκριναν διαφορετικά πρότυπα συμπίεσης και πολλά από τα συμπεράσματα των αποτελεσμάτων τους είναι εντελώς διαφορετικά. Κάποια από τα συμπεράσματα είναι ασαφή και δεν είναι ξεκάθαρο ποιος κωδικοποιητής κυριαρχεί όσο αφορά την ποιότητα και την αποδοτικότητα. Οι συγκρίσεις που ακολουθούν αφορούν το σύνολο δεδομένων με βίντεο από το Netflix, όμως συνάδουν και με τα αποτελέσματα από το σύνολο δεδομένων από σκηνικά έκτασης ανάγκης, καθώς οι μετρήσεις στα δύο σύνολα συμφωνούν.

Στο άρθρο [38], αναφέρεται ότι η απόδοση κωδικοποίησης για το AV1, είναι χαμηλότερη από τους κωδικοποιητές x264 και HM, καθώς επιτυγχάνεται αύξηση του bitrate έως 10.5% και 67.7% αντίστοιχα για την ίδια ποιότητα βίντεο. Συγκρίνοντας το AV1 με το x265, το AV1 έχει overhead 39.1% και η κωδικοποίηση είναι πάνω από δύο φορές πιο αργή, ενώ ο HM έχει κέρδος κωδικοποίησης 13.9% σε σχέση με το x265, ενώ η κωδικοποίηση είναι τρεις φορές πιο αργή. Αυτά αποτελέσματα συμφωνούν είναι πιο κοντά στα δικά μας όταν έχουμε βίντεο με 50fps, καθώς το HM μειώνει τις απαιτήσεις κατά 15.37%, ενώ για βίντεο με 25fps, έχει κέρδος μόνο 8.9%.

Στο άρθρο [39] αναφέρεται ότι το AV1, έχει κέρδος σε bitrate, κατά 17% στο HM, και κατά 13% στο πρότυπο VP9, ενώ στο άρθρο[40] αναφέρεται εξοικονόμηση του AV1 έως και 17% σε σχέση με το VP9. Από την άλλη το Facebook [41] αναφέρει ότι το AV1 ξεπέρασε το VP9 κατά 30%. Στα δικά μας αποτελέσματα, για βίντεο 50fps, σε σύγκριση με το VP9, το AV1 είναι περίπου κατά 15% πιο αποτελεσματικό από ότι αναφέρεται στο [39], αφού πετυχαίνει κέρδος 28%. Όσο αφορά το HM, το 17% συμφωνεί με το 16.8% των αποτελεσμάτων μας στα βίντεο με 50fps, ενώ για 25fps υπάρχει μείωση στο 10.25%.

Μια μελέτη από το Netflix [42] σύγκρινε τους κωδικοποιητές με video on demand adaptive streaming. Φαίνεται ότι όλοι οι κωδικοποιητές x265, HM, VP9 και AV1 παρουσιάζουν καλύτερη απόδοση από το πρότυπο x264. Όπως παρουσιάζονται τα αποτελέσματα τον πίνακα σύγκρισης BD-Rate του PSNR, το AV1 παρουσιάζει καλύτερα αποτελέσματα από το x265 και το VP9, αφού μειώνει την απόδοση κατά 15.21% και 18.70% αντίστοιχα. Το x265 είναι καλύτερο από το VP9 κατά 4.11%. Άρα η απόδοση των προτύπων κωδικοποίησης VP9 και x265 είναι πολύ κοντά. Γενικά παρατηρούμε ότι η σύγκριση του VP9 και του x265 δεν είναι σταθερή, καθώς στα αποτελέσματα μας για βίντεο με 25fps, το x265 έχει καλύτερη απόδοση από το VP9 κατά 11.92%, ενώ για 50fps φαίνεται το VP9 είναι καλύτερο από το x265 κατά 4.732%.

Σε μια πρόσφατη μελέτη [21], συγκρίνονται βίντεο ανάλυσης 768x432p, όπως και στην παρούσα μελέτη και παρατηρούμε ότι επιβεβαιώνεται το συμπέρασμα ότι το πρότυπο κωδικοποίησης VVC, είναι καλύτερο από το x265, AV1 και VP9. Συγκεκριμένα το VVC σε σχέση με το AV1, το x265 και το VP9, υπερτερεί κατά 56.17%, 67.50% και 75% αντίστοιχα.

Το AV1 εξοικονομεί περίπου 23.33% έναντι του x265 και 36.52% έναντι του VP9. Το x265 αποδίδει 25.30%, κέρδος bitrate, έναντι του VP9. Όσο αφορά τα βίντεο από το σύνολο του Netflix, το AV1 είναι καλύτερο από το x265 κατά 18.654% και από το VP9 κατά 28.248%. Άρα τα αποτελέσματα σε αυτή την περίπτωση συμφωνούν.

Όσο αφορά την παρούσα μελέτη το VVC για βίντεο με 25fps υπερτερεί από το AV1 κατά 26.948%, από το x265 κατά 40.24% , και από το VP9 κατά 46.499%, ενώ για βίντεο για 50fps, υπερτερεί από το AV1 κατά 22.948%, από το x265 κατά 34.266% , και από το VP9 κατά 43.084%. Παρατηρούμε ότι δεν έχουμε πολύ μεγάλη απόκλιση στο ποσοστό του BD-Rate για το VVC, σε σχέση με την διαφορά στα frame per second των βίντεο.

Κεφάλαιο 4

Προσαρμοστική Μετάδοση Βίντεο

4.1	Πρωτόκολλα Επικοινωνίας	49
	4.1.1 Progressive Video Streaming	50
	4.1.2 Adaptive Video Streaming	51
4.2	MPEG-DASH	53
	4.2.1 Περιγραφή MPEG-DASH	54
	4.2.2 Media Presentation Description (MPD)	57
	4.2.3 Μορφή τμημάτων- Segments	60
4.3	Περιγραφή Συστήματος Προσαρμοστικής Μετάδοσης	61

Σε αυτό το κεφάλαιο θα αναλυθεί εκτενώς το MPEG-DASH, ένα πρότυπο προσαρμοστικής μετάδοσης (Adaptive streaming), που βασίζεται στο πρωτόκολλο HTTP. Η βασική ιδέα είναι ότι το βίντεο χωρίζεται σε μικρά κομμάτια που το κάθε ένα περιέχει ένα μικρό μέρος του συνολικού βίντεο. Υπάρχουν ποικίλα διαθέσιμα bitrates από τα οποία ο MPEG-DASH Client, επιλέγει αυτόματα το επόμενο κομμάτι περιεχομένου που θα κατεβάσει και θα παίξει με βάση τις παρούσες συνθήκες του δικτύου. Με αυτό τον τρόπο προσφέρεται στον χρήστη το υψηλότερο δυνατό bitrate, χωρίς να υπάρχουν διακοπές, καθυστερήσεις ή rebuffering. Πιο κάτω παρουσιάζεται εκτενώς, η διαδικασία [10] και το σύστημα προσαρμοστικής μετάδοσης το οποίο υλοποιήθηκε.

4.1 Πρωτόκολλα Διάδοσης Βίντεο

Η μετάδοση βίντεο μέσω του Διαδικτύου ξεκίνησε τη δεκαετία του 1990 και έφερε στην επιφάνεια προκλήσεις όπως η παράδοσή του χωρίς μεγάλη καθυστέρηση και χωρίς κατανάλωση μεγάλου όγκου δεδομένων. Το πρωτόκολλο μεταφοράς δεδομένων σε πραγματικό χρόνο ονομάζεται RTP. Μέσω του πρωτοκόλλου αυτού καθορίζεται η μορφή που πρέπει να έχουν πακέτα βίντεο και ήχου, ώστε να μεταδίδονται με χαμηλό overhead [43]. Ένα RTP πακέτο περιέχει:

- Payload type identification (είδος δεδομένων, π.χ. ήχος ή βίντεο, πρότυπο κωδικοποίησης, μέγεθος, αποστολέα)
- Packet sequence numbering (αρίθμηση ακολουθίας πακέτων) και
- Timestamping (χρονοσήμανση) [12]

Το RTP λειτουργεί καλά σε διαχειριζόμενα δίκτυα IP. Ωστόσο, σήμερα, τα διαχειριζόμενα δίκτυα έχουν αντικατασταθεί από δίκτυα μετάδοσης περιεχομένου (CDN), πολλά από τα οποία δεν υποστηρίζουν ροή RTP. Το πρωτόκολλο RTP απαιτεί από το διακομιστή να διαχειρίζεται μια ξεχωριστή ροή για κάθε πελάτη και έτσι υπάρχει μεγάλη κατανάλωση πόρων. Υπήρξε ανάγκη για την δημιουργία αυτού του πρωτοκόλλου διότι τα δύο γνωστά πρωτόκολλα γενικής χρήσης TCP και UDP δεν μπορούν να χρησιμοποιηθούν για τέτοιες εφαρμογές. Το User Datagram Protocol (UDP) είναι ένα πρωτόκολλο το οποίο δεν παρέχει καμία επιβεβαίωση (acknowledgment) προς το αποστολέα και δεν εγγυάται αξιόπιστη επικοινωνία. Τα πακέτα UDP που αποστέλλονται από τον αποστολέα, μπορεί να μην φτάσουν καθόλου στον παραλήπτη σε περίπτωση που το δίκτυο είναι υπερφορτωμένο. Υπάρχει επίσης η πιθανότητα τα πακέτα να φτάσουν με λάθος σειρά ή ακόμα και να φτάσουν δυο φορές το καθένα [44]. Από την άλλη το Transmission Control Protocol (TCP) είναι ένα αξιόπιστο πρωτόκολλο σύνδεσης που αποστέλλει στον αποστολέα επιβεβαίωση (acknowledgment). Η αξιοπιστία του έγκειται στο γεγονός ότι αν κάποιο πακέτο χαθεί κατά την αποστολή του τότε είναι υπεύθυνο να το ξανά στείλει και επίσης αν κάποιο από τα πακέτα στάλθηκε περισσότερο από μια φορές, τότε είναι υπεύθυνο να απορρίψει τις πλεονάζον αποστολές του. Το RTP δεν καταγράφει την δέσμευση των πόρων και δεν εγγυάται ποιότητα υπηρεσίας όσο αφορά τις πραγματικού χρόνου υπηρεσίες. RTP δεν παρέχει οποιουδήποτε μηχανισμούς που να εγγυούνται την ποιότητα της υπηρεσίας όπως η έγκαιρη παράδοση των πακέτων

Τα RTP πακέτα ενσωματώνονται σε UDP τμήματα, όμως η ενσωμάτωση αυτή είναι ορατή μόνο end-to-end, δεν είναι ορατή από ενδιάμεσους διακομιστές[12]

Το πρωτόκολλο the Real Time Control Protocol (RTCP) αποτελεί μέρος του RTP και είναι υπεύθυνο για τον έλεγχο της ποιότητας της υπηρεσίας (QoS) και για να βελτιώνει την μεταφορά δεδομένων, κάτι που δεν παρέχει το RTP. Αυτό το πρωτόκολλο ελέγχου, παρακολουθεί την παράδοση των δεδομένων ανάλογα με το μέγεθος του multicast δικτύου, και προσφέρει κάποιο μικρό ποσοστό ελέγχου καθώς και λειτουργία αναγνώρισης.

Σε ένα RTP session κάθε συμμετέχοντας αποστέλλει περιοδικά RTCP πακέτα ελέγχου στους υπόλοιπους συμμετέχοντες. Σε κάθε RTCP πακέτο περιέχονται στατιστικές αναφορές τις οποίες στέλνουν οι παραλήπτες και ο αποστολέας και αφορούν την ποιότητα υπηρεσίας, τον συνολικό αριθμό πακέτων που στάλθηκαν, των πακέτων που χάθηκαν, την καθυστέρηση παραλαβής μεταξύ πακέτων και άλλα. Αυτή η ανατροφοδότηση μπορεί να χρησιμοποιηθεί για τον έλεγχο ποιότητας και απόδοσης και ο αποστολέας μπορεί να προσαρμόσει τον ρυθμό μετάδοσής του ανάλογα [12]. Τα RTP και RTCP πακέτα διαχωρίζονται μεταξύ τους χρησιμοποιώντας διαφορετικά port numbers, ενώ για την μετάδοση των πακέτων χρησιμοποιούν την ίδια διεύθυνση δικτύου IP multicast address η οποία ανταποκρίνεται στην συγκεκριμένη σύνοδο RTP [8].

Στο παρελθόν οι περισσότερες τεχνολογίες για μεταδόσεις βίντεο χρησιμοποιούσαν πρωτόκολλα μεταδόσεις όπως το RTP, όπου υπήρχε progressive video streaming.

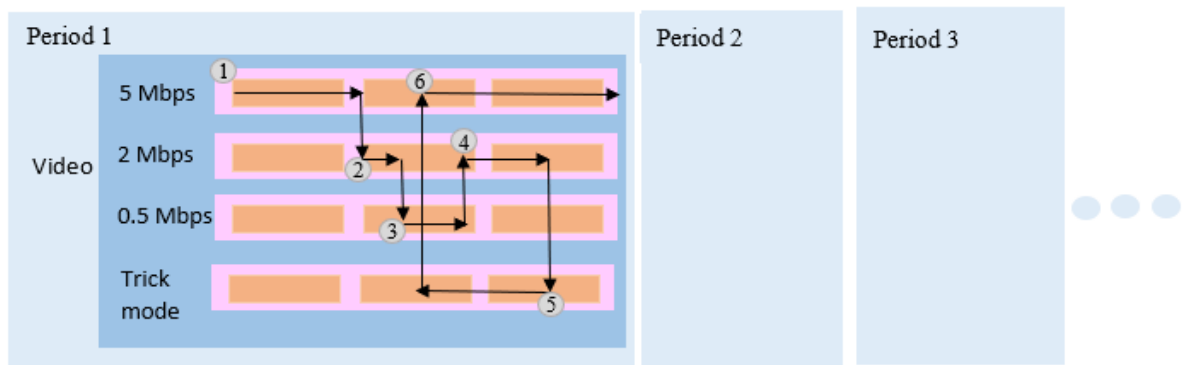
4.1.1 Progressive Video Streaming

Αναφέρεται σε ένα ενιαίο βίντεο που μεταδίδεται μέσω του Διαδικτύου. Το ίδιο βίντεο μπορεί να προσαρμοστεί σε διάφορα μεγέθη οθόνης, ανεξάρτητα από την συσκευή στην οποία αναπαράγεται. Ωστόσο κατά την αναπαραγωγή συνήθως προκύπτουν προβλήματα στην ποιότητα αλλά και το φόρτωμα του βίντεο (buffering). Για παράδειγμα ένα βίντεο 1280x720 δεν θα έχει την ίδια ποιότητα αν αναπαραχθεί σε μεγαλύτερη οθόνη π.χ. 1920x1080 pixel. Προβλήματα buffering δημιουργούνται συνήθως εάν υπάρχει κακή ή αδύνατη σύνδεση στο διαδίκτυο και η ροή του βίντεο (video stream), δεν μπορεί να κατεβεί γρήγορα. Σε αυτή την περίπτωση το βίντεο πρέπει να σταματήσει, να περιμένει περισσότερα δεδομένα και στην συνέχεια να ξεκινήσει ξανά [43].

4.1.2 Adaptive Video Streaming

Τα παραπάνω έρχονται σε αντίθεση με την προσαρμοστική μετάδοση (Adaptive) που χρησιμοποιούν οι σημερινές τεχνολογίες και βασίζεται σχεδόν αποκλειστικά στο πρωτόκολλα HTTP. Το HTTP έχει πολλά οφέλη, καθώς μπορεί να υποστηρίζεται από το CDN, και να είναι φιλικό προς το τείχος προστασίας, σε αντίθεση με το RTP. Επίσης με το HTTP ο πελάτης διαχειρίζεται τη ροή χωρίς να χρειάζεται να διατηρήσει ένα session στον διακομιστή. Επομένως, η παροχή μεγάλου αριθμού πελατών ροής δεν επιβάλλει επιπλέον κόστος σε πόρους διακομιστή πέραν της τυπικής χρήσης HTTP από τον Ιστό [43].

Το Adaptive Video Streaming λύνει τα προβλήματα του Progressive Video Streaming. Είναι μια τεχνολογία που σχεδιάστηκε για να παρέχει βίντεο στον χρήστη με αποτελεσματικότητα και με την υψηλότερη δυνατή ποιότητα για κάθε συγκεκριμένο χρήστη. Μπορεί να δώσει λύση στο πρόβλημα της ποιότητας του βίντεο επιτρέποντας στον video provider να δημιουργήσει ένα διαφορετικό βίντεο για κάθε μέγεθος οθόνης ή συσκευής. Με αυτό τον τρόπο ο χρήστης βλέπει πάντα καλή ποιότητα βίντεο. Το φαινόμενο του buffering εμφανίζεται όταν ο χρήστης δεν μπορεί να κατεβάσει ένα βίντεο αρκετά γρήγορα για αναπαραχθεί. Τα περισσότερα βίντεο χρειάζονται 24 frames per second για να παίξουν, άρα η σύνδεση πρέπει να κατεβάζει τουλάχιστον 24 frames per second για να αποφευχθεί το buffering. Με το Adaptive Video Streaming προσαρμόζεται η ταχύτητα σύνδεσης στο διαδίκτυο του χρήστη. Η προσαρμογή αυτή (Adaptation), καθορίζει ποιο βίντεο είναι καλύτερο για κάθε χρήστη και οι ρυθμίσεις αυτές αλλάζουν από δευτερόλεπτο σε δευτερόλεπτο. Καθώς αλλάζει η σύνδεση των χρηστών στο διαδίκτυο, η ροή προσαρμοστικότητας (Adaptive stream) αλλάζει μεταξύ των ποιοτήτων του βίντεο και αυτό επιτυγχάνεται με την χρήση των τμημάτων (segments) [47].



Σχήμα 4.1 : Ένα παράδειγμα δυναμικής προσαρμοστικής ροής βίντεο. Οι αριθμημένοι κύκλοι (label) δείχνουν τα σημεία δράσης που λαμβάνει η συσκευή

Στο σχήμα 4.1 περιγράφεται ένα παράδειγμα απλής περίπτωσης χρήσης δυναμικής ροής περιεχομένου πολυμέσων και συγκεκριμένα ενός βίντεο (Adaptive video streaming). Γενικά το περιεχόμενο μπορεί να αποτελείται από στοιχεία βίντεο και ήχου, εμείς θα ασχοληθούμε μόνο με το βίντεο, το οποίο μελετήθηκε και στην παρούσα εργασία.

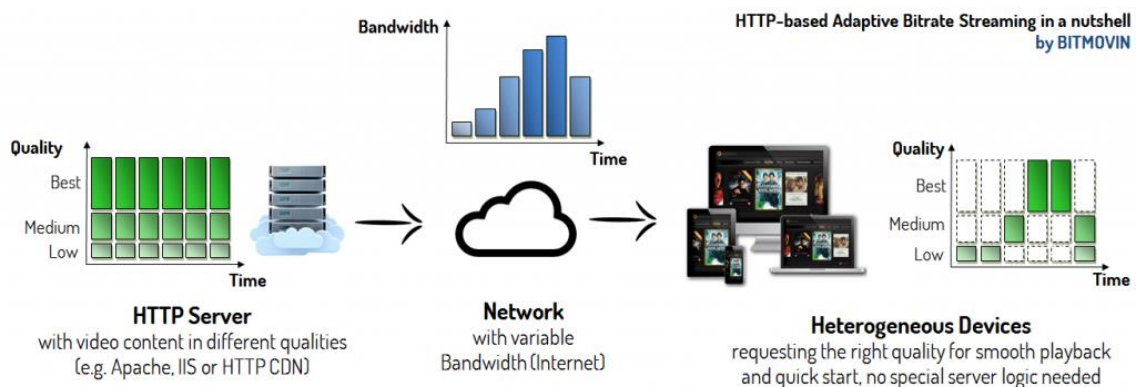
Η πηγή του βίντεο κωδικοποιείται σε τρία διαφορετικά bitrates: 5 Mbytes, 2 Mbytes και 500 kilobits per second (Kbps). Κατά την διάρκεια αναπαραγωγής παρέχεται, ένα I-frame που έχει χαμηλό frame rate.

Υποθέτουμε ότι η συσκευή ξεκινά τη ροή του περιεχομένου ζητώντας τμήματά από το βίντεο bitstream, με την μεγαλύτερη διαθέσιμη ποιότητα 5 Mbps (label 1). Αφού μεταδοθεί το πρώτο τμήμα του βίντεο, παρακολουθείται το διαθέσιμο εύρος ζώνης, και παρατηρείται ότι το πραγματικό διαθέσιμο εύρος ζώνης είναι μικρότερο από 5 Mbps. Αρά το επόμενο διαθέσιμο σημείο μεταγωγής, μετατρέπει το βίντεο σε ταχύτητα 2 Mbps (label 2). Η συσκευή συνεχίζει να παρακολουθεί το πραγματικό εύρος ζώνης του διαδικτύου και παρατηρείται ότι αυτό μειώθηκε σε τιμή μικρότερη από 2 Mbps. Επομένως, για να διατηρηθεί η συνεχής αναπαραγωγή του βίντεο, χωρίς να κολλάει, η συσκευή μειώνει την ροή του βίντεο σε 500 kbps (label 3). Το περιεχόμενο συνεχίζει να αναπαράγεται με αυτές τις τιμές, μέχρι το εύρος ζώνης να αυξηθεί. Όταν γίνει αυτό το βίντεο μεταφέρεται στα 2 MB (label 4). Στην συνέχεια, ο χρήστης αποφασίζει να κάνει pause και rewind, δηλαδή να σταματήσει και να επιστρέψει προς τα πίσω. Σε αυτή την περίπτωση, η συσκευή θα ξαναρχίσει την ροή του βίντεο από την μεγαλύτερη ποιότητα 5Mb. (label 6).

4.2 MPEG-DASH

Το MPEG DASH (Dynamic Adaptive Streaming over HTTP) είναι ένα πρότυπο που στοχεύει στην παροχή συνεχούς ροής βίντεο σε διαφορετικές συσκευές και σε περιπτώσεις όπου οι συνθήκες του διαδικτύου μεταβάλλονται. Πολλοί οργανισμοί αναπτύσσουν ποικίλες τεχνολογίες για την μεταφορά πολυμέσων στο διαδίκτυο και την μετάδοσή τους σε κινητές συσκευές. Το MPEG-DASH έχει τυποποιηθεί από το 2009 και έγινε πρότυπο ISO / IEC (International Electrotechnical Commission) το 2012 επιδιώκοντας την προσαρμοστική ροή δεδομένων. Στη συνέχεια, το DASH υιοθετήθηκε από άλλα τυποποιημένα συστήματα πολυμέσων streaming όπως το IPTV (Open IPTV standard) και το Digital Video Broadcasting (DVB) ως πρότυπο για τη μεταφορά βίντεο μέσω του Διαδικτύου. Χρησιμοποιεί τον αλγόριθμο Adaptive Bitrate (ABR), ο οποίος έχει ως πρώτο στόχο να μεγιστοποιεί την αντιληπτή ποιότητα της εμπειρίας (QoE) του χρήστη αφού προσαρμόζεται στις συνθήκες δικτύου που μεταβάλλονται συνεχώς [46].

Το DASH Industry Forum (DASH-IF) προωθεί την υιοθέτηση και την έρευνα σύγχρονων τεχνολογιών προσαρμοστικής ροής. Το Dash-IF προσφέρει πλήρη λειτουργικότητα και συγκεκριμένες οδηγίες υλοποίησης, μέσω του documentation που παρέχει. Η κοινότητα έχει επίσης αναπτύξει έναν ανοικτού κώδικα dash.js reference player, ο οποίος χρησιμοποιείται για αναπαραγωγή βίντεο με στόχο την δοκιμή και την έρευνα τους. Επίσης, το DASH-IF διαθέτει στο κοινό ένα ολοκληρωμένο σύνολο δεδομένων που μπορούν να χρησιμοποιηθούν για δοκιμές, καθώς επίσης και video players/clients και λογισμικό για την προετοιμασία και την επικύρωση του περιεχομένου των βίντεο [47].



Σχήμα 4.2: HTTP Adaptive Video Streaming

4.2.1 Περιγραφή MPEG-DASH

Ένα αρχείο πολυμέσων αποθηκεύεται σε ένα διακομιστή HTTP. Το περιεχόμενο υπάρχει στον διακομιστή σε δύο μέρη, στο Media Presentation Description (MPD) και τα τμήματα, segments. Το Media Presentation Description (MPD) είναι ένα αρχείο, manifest το οποίο παρέχει πλήρη στοιχεία όσο αφορά το διαθέσιμο περιεχόμενο, τις ποικίλες εναλλακτικές λύσεις, τις URL διευθύνσεις του βίντεο και άλλα. Τα τμήματα περιλαμβάνουν το περιεχόμενο του βίντεο μοιρασμένο σε «κομμάτια» σε ένα ή περισσότερα αρχεία. Για να αναπαραχθεί το περιεχόμενο του βίντεο ο DASH client ζητά πρώτα το MPD και μετά με βάση αυτές τις πληροφορίες ζητά τα τμήματα που ταιριάζουν καλύτερα στις απαιτήσεις του. Το MPD μπορεί να παραδοθεί χρησιμοποιώντας HTTP, ηλεκτρονικό ταχυδρομείο, broadcast ή άλλες μεταφορές. Μέσω του MPD, ο χρήστης μαθαίνει για την διαθεσιμότητα του μέσου, τους τύπους μεσών, τις χωρικές αναλύσεις, το ελάχιστο και το μέγιστο εύρος ζώνης και άλλα χαρακτηριστικά. Χρησιμοποιώντας αυτές τις πληροφορίες ο DASH client επιλέγει την κατάλληλη κωδικοποίηση και ξεκινά η αναπαραγωγή του περιεχομένου, μεταφέροντας τα τμήματα με HTTP GET requests. Αφού γίνει το κατάλληλο buffering, ώστε να επιτρέπονται οι αλλαγές στην απόδοση του δικτύου, ο πελάτης συνεχίζει να ανακτά τα επόμενα τμήματα και να παρακολουθεί τις διακυμάνσεις του εύρου ζώνης. Ανάλογα με τις μετρήσεις που παίρνει αποφασίζει αν θα επιλέγει τα τμήματα με μικρότερο ή υψηλότερο bitrate, έτσι ώστε να προσαρμόζετε στο διαθέσιμο εύρος ζώνης και να διατηρεί ένα επαρκές buffer. Το MPEG-DASH καθορίζει μόνο το MPD και την μορφή των τμημάτων. Ο τρόπος που παραδίνονται τα MPD, οι μορφές που περιέχουν τα τμήματα, η συμπεριφορά του πελάτη και η αναπαραγωγή του περιεχομένου δεν εμπίπτουν στο πεδίο εφαρμογής του MPEG-DASH.

Μια από τις πολλές προκλήσεις που έχουν να αντιμετωπίσουν οι πάροχοι μετάδοσης βίντεο, είναι να παρέχουν την ποιότητα που επιθυμεί ο χρήστης, ανταποκρινόμενοι στους περιορισμένους πόρους του δικτύου. Για να αντιμετωπίσουν αυτήν την πρόκληση, πρέπει να μελετήσουν εκατοντάδες μορφές κωδικοποίησης και σχετικές απαιτήσεις αποθήκευσης, προκειμένου να βελτιστοποιηθεί η ποιότητα μετάδοσης βίντεο κατ' απαίτηση (On demand) ή σε πραγματικό χρόνο (live) [21]. Δημοφιλείς πάροχοι όπως το YouTube, το Netflix και το Amazon επιλύουν αυτό το πρόβλημα ποιότητας μετάδοσης σε πραγματικό χρόνο αποθηκεύοντας εκατοντάδες προ-κωδικοποιημένες μορφές βίντεο

και παραδίδοντας τις με βάση τις ανάγκες των χρηστών. Για παράδειγμα, το YouTube χρησιμοποιεί ένα Νευρωνικό [48] δίκτυο για να προσφέρει προσαρμοστική μετάδοση bitrate (ABR) στο διαδίκτυο. Λόγω του ότι εκατομμύρια βίντεο παρακολουθούνται καθημερινά [49] το YouTube χρησιμοποιεί πολλαπλές κωδικοποιήσεις βίντεο, με διαφορετικά bitrate η κάθε μια [6] για κάθε τμήμα ABR, έτσι δεν απαιτείται να γίνεται κωδικοποίηση πολλές φορές για να αναπαραχθούν πολλά βίντεο στους χρήστες. Το Νευρωνικό δίκτυο που χρησιμοποιείται μαθαίνει με βάση το περιεχόμενο του βίντεο και ενημερώνει τις παραμέτρους του δικτύου χρησιμοποιώντας απλές λειτουργίες που

Type	Video Bitrate, Standard Frame Rate (24, 25, 30)	Video Bitrate, High Frame Rate (48, 50, 60)
2160p (4k)	35-45 Mbps	53-68 Mbps
1440p (2k)	16 Mbps	24 Mbps
1080p	8 Mbps	12 Mbps
720p	5 Mbps	7.5 Mbps
480p	2.5 Mbps	4 Mbps
360p	1 Mbps	1.5 Mbps

λαμβάνονται από τα τμήματα που είναι διαχωρισμένο το βίντεο. Ουσιαστικά το YouTube, όπως και άλλοι παροχείς βίντεο, χρησιμοποιούν το MPEG-DASH, ώστε να πάρουν την ανάλυση των βίντεο, ανάλογα με το framerate. Στην συνέχεια ανάλογα με το διαθέσιμο εύρος ζώνης, γίνεται προεπιλογή του αντίστοιχου bitrate. Ο στόχος είναι η μεγιστοποίηση της ποιότητας του βίντεο ενώ συμμορφώνεται με τους περιορισμούς του εύρου ζώνης, όπως προτείνονται από το YouTube [6]. Στο σχήμα 4.2.1 παρουσιάζονται οι τιμές των bitrate που προτείνει το YouTube για διάφορες αναλύσεις βίντεο [50]. Μέσα από μια μελέτη χρησιμοποιώντας βίντεο που περιλαμβάνουν πολλές κινήσεις, σύμφωνα με τις προτάσεις του YouTube, για βίντεο με ανάλυση 1080p, το συνιστώμενο bitrate είναι 12000kbps.

Σχήμα 4.2.1: YouTube Recommended Bitrates for different resolutions[6]

Υπάρχουν πολλά οφέλη και αρκετοί λόγοι, όπου μεγάλα ονόματα εταιρειών, όπως το YouTube, και το Netflix στρέφονται προς το MPEG-DASH, μερικοί από αυτούς τους λόγους παρουσιάζονται πιο κάτω.

Πλεονεκτήματα MPEG DASH

- Το MPEG-DASH είναι μια ανεξάρτητη, ανοιχτή και διεθνής εταιρεία, η οποία έχει ευρεία υποστήριξη από τον κλάδο, όπως φαίνεται από τη μεγάλη λίστα μελών του DASH Industry Forum [52].
- Το MPEG-DASH μπορεί να επιτρέπεται πλέον σε πολλά νέα πρότυπα, όπως η τυποποιημένη αναπαραγωγή σε πλατφόρμες HTML5. Οπότε το MPEG-DASH υποστηρίζεται και από τους περισσότερους browsers.
- Οι πελάτες χρησιμοποιούν το καθιερωμένο πρωτόκολλο HTTP, μέσω του οποίου παρέχεται πρόσβαση παντού, αφού μπορεί να διασχίσει firewalls, αντίθεση με το TCP.
- Οι διακομιστές του DASH, είναι κεντρικοί Web servers, οι οποίοι μειώνουν κατά πολύ το λειτουργικό κόσμος και επιτρέπουν την ανάπτυξη κρυφής μνήμης που βοηθούν στην βελτίωση της απόδοσης.
- Ο πελάτης ζητά κάθε κομμάτι του βίντεο ανεξάρτητα και διατηρεί την κατάσταση αναπαραγωγής σταθερή. Αυτό σημαίνει ότι ο πελάτης μπορεί να ανακτήσει κομμάτια βίντεο και από διαφορετικούς διακομιστές με χαμηλή ανοχή σφάλματος.
- Αυτά τα πλεονεκτήματα επιτρέπουν στους παρόχους υπηρεσιών να αξιοποιήσουν τις υφιστάμενες και σημαντικά φθηνότερες υποδομές HTTP. Κατοχυρωμένα εμπορικά συστήματα, όπως Microsoft's Smooth Streaming, Adobe's HTTP Dynamic Streaming (HDS) και Apple's HTTP Live Streaming (HLS), εκμεταλλεύονται τα υπάρχοντα CDNs και τις προσωρινές κρυφές μνήμες [43].

Χαρακτηριστικά MPEG DASH

Κάποια καθιερωμένα χαρακτηριστικά του MPEG DASH:[52]

- Το MPD παρέχει αρκετές πληροφορίες στον πελάτη έτσι ώστε να μπορεί να επιλέξει και να εναλλάξει ροές (streams). Για παράδειγμα, μπορεί να επιλέξει διαφορετικές γωνιές της κάμερας για ένα βίντεο, ή να επιλέξει υπότιτλους μεταξύ διαφορετικών γλωσσών. Κάνοντας τις επιλογές αυτές δυναμικά μεταβάλλεται και το bitrate του βίντεο.
- Ανάμεσα στις περιόδους και τα τμήματα μπορεί να εισαχθούν διαφημίσεις, ακόμα και σε ζωντανές περιπτώσεις.

- Οι URL διευθύνσεις των τμημάτων μπορούν να σηματοδοτηθούν χρησιμοποιώντας ένα σχέδιο προτύπου με βάση το MPD. (Compact manifest)
- Το MPD μπορεί να χωριστεί σε πολλά κομμάτια και κάποια από αυτά μπορούν να αναφερθούν και εξωτερικά, επιτρέποντας με αυτό τον τρόπο την λήψη του MPD σε πολλά στάδια. (Fragmented manifest.)
- Η διάρκεια του κάθε τμήματος μπορεί να διαφέρει. Στην περίπτωση της ζωντανής μετάδοσης η διάρκεια του επόμενου τμήματος μπορεί να σηματοδοτηθεί με την παράδοση του τρέχοντος τμήματος.
- Το ίδιο περιεχόμενο μπορεί να είναι διαθέσιμο σε πολλές διευθύνσεις URLs, δηλαδή σε διαφορετικούς διακομιστές ή CDNs. Ως εκ τούτου ο πελάτης μπορεί να μεταδώσει οποιοδήποτε από αυτά, αξιοποιώντας το μέγιστο διαθέσιμο εύρος ζώνης του δικτύου.

4.2.2 Media Presentation Description (MPD)

Το Multimedia/Media Presentation Description είναι ένα XML αρχείο που παρουσιάζει τις διάφορες ποιότητες του βίντεο και των επιμέρους τμημάτων για κάθε ποιότητα με HTTP URLs. Περιέχει μια πλήρης περιγραφή των προσβάσιμων τμημάτων του βίντεο και αποτελεί ένα ιεραρχικό μοντέλο δεδομένων το οποίο αποτελείται από μια ή περισσότερες περιόδους.

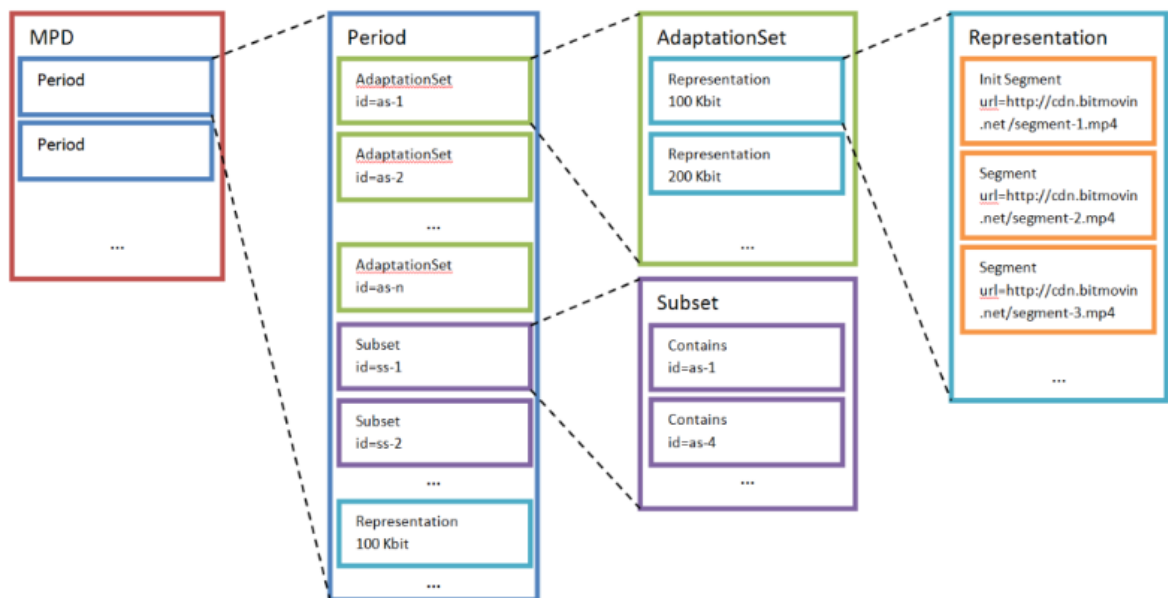
Κάθε περίοδος έχει χρόνο έναρξης και διάρκεια και συμπεριλαμβάνει media components, όπως για παράδειγμα βίντεο διαφορετικών κωδικοποιητών, ήχων, υποτίτλων. Τα στοιχεία αυτά έχουν κάποια χαρακτηριστικά τα οποία δεν αλλάζουν κατά την διάρκεια μιας περιόδου όπως το bitrate και το frame per seconds. Τα media components είναι διατεταγμένα σε ένα σει προσαρμογής. Οπότε, κάθε περίοδος (period) μπορεί να περιέχει ένα ή περισσότερα σύνολα προσαρμογής (adaptation set) [53].

Το κάθε σύνολο ενσωματώνει μια ή περισσότερες κωδικοποιήσεις ψηφιακού περιεχομένου. Με αυτό τον τρόπο μπορεί να γίνει ομαδοποίηση των διαφόρων στοιχείων που ανήκουν στην ίδια κατηγορία. Κάθε adaptation set περιλαμβάνει πληροφορίες σχετικά με ένα ή περισσότερα συστατικά των πολυμέσων και τις ποικίλες κωδικοποιημένες εναλλακτικές λύσεις. Δηλαδή στο ίδιο adaptation set μπορούν να βρίσκονται στοιχεία με την ίδια ποιότητα, ήχο, γλώσσα, κωδικοποιητή κτλ. Για

παράδειγμα ένα adaptation set μπορεί να περιέχει τα διαφορετικά bitrate για τα συστατικά του βίντεο του ίδιου πολυμέσου, και άλλο adaptation set, μπορεί να περιέχει διαφορετικά bitrate, όσο αφορά τον ήχο του ίδιου πολυμέσου. Συνήθως, ένα Adaptation set αποτελείται από ένα σύνολο αναπαραστάσεων (representations) που περιλαμβάνουν διαφορετικά resolutions,bitrates κτλ.

Μια αναπαράσταση αποτελείται από ένα ή περισσότερα τμήματα. Αυτά τα τμήματα είναι τα «κομμάτια» ροής των πολυμέσων σε μια χρονική ακολουθία. Κάθε τμήμα περιγράφεται με ένα URI, δηλαδή μια διεύθυνση στον διακομιστή, που μπορεί να κατεβεί χρησιμοποιώντας HTTP GET. Τα τμήματα εκτός από το URI, κάποτε περιγράφονται και με ένα επιπρόσθετο εύρος byte στην περίπτωση που είναι αποθηκευμένα σε ένα μεγαλύτερο συνεχές αρχείο. Μια αναπαράσταση είναι ουσιαστικά μια κωδικοποιημένη εναλλακτική του ίδιου συστατικού. Για παράδειγμα, η αναπαράσταση αυτή διαφέρει από τις άλλες ως προς το εύρος ζώνης και το resolution [54].

Ο DASH client, πρώτα αναλύει το MPD αρχείο και με βάση τα περιεχόμενα, τις δυνατότητες του και τις επιλογές του χρήστη, επιλέγει το σύνολο των αναπαραστάσεων που θα χρησιμοποιήσει. Στην συνέχεια, δημιουργεί ένα χρονικό πλαίσιο όπου αρχίζει να αναπαράγεται το περιεχόμενο του πολυμέσου ζητώντας κατάλληλα τμήματα.



Σχήμα 4.2.2: Media Presentation Description (MPD): Περιγράφει τα προσβάσιμα τμήματα του βίντεο και τους αντίστοιχους χρονισμούς [46]

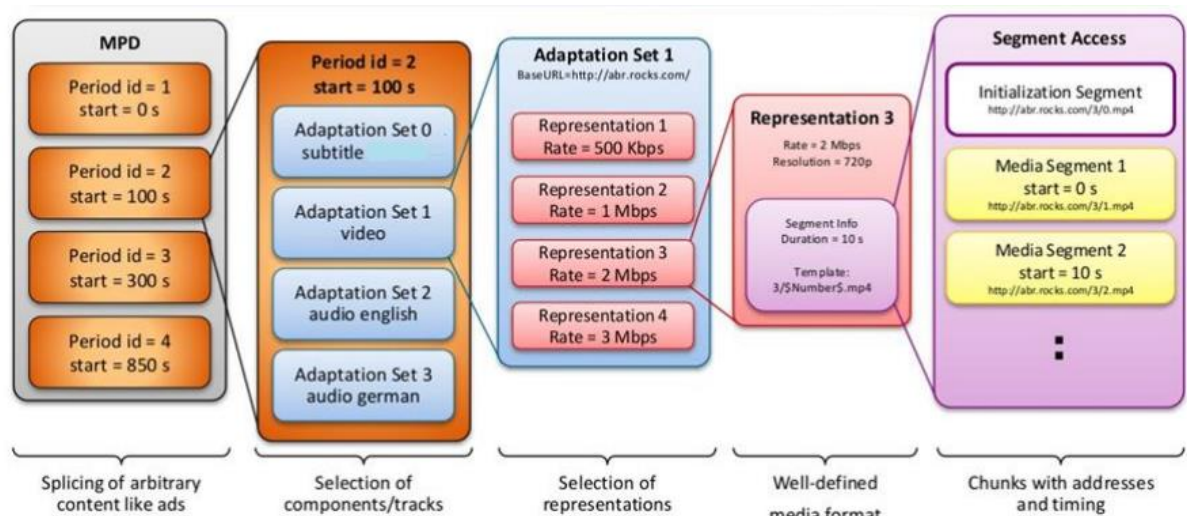
Η μετάδοση του βίντεο μέσω HTTP υλοποιείται στο επίπεδο των αναπαραστάσεων που κατεβάζει ο παραλήπτης. Μέσω των περιόδων και των αναπαραστάσεων επιτυγχάνεται ο συγχρονισμός και ομαλή μετάβαση και προσαρμογή από ένα τμήμα σε κάποιο άλλο μεταξύ μιας ή περισσότερων αναπαραστάσεων. Η διαδικασία της προσαρμογής πραγματοποιείται στο επίπεδο του παραλήπτη, ο οποίος κατεβάζει το αρχείο MPD, και στην συνέχεια κατεβάζει διαδοχικά τα τμήματα του βίντεο, βάση του playout buffer και του διαθέσιμου εύρους ζώνης το οποίο υπολογίζεται τοπικά. Οι επιλογές που καθορίζουν ποια αναπαράσταση θα επιλεγεί για να αναπαραχθεί καθορίζονται από την κατάσταση το buffer, τις συνθήκες του δικτύου, πόσο γρήγορο ή αργό είναι για παράδειγμα το διαδίκτυο. Επίσης η επιλογή εξαρτάται από τους διαθέσιμους πόρους και άλλες δραστηριότητες οι οποίες εκτελούνται στην συσκευή, καθώς και από την ανάλυση βίντεο την οποία θα επιλέξει ο χρήστης, λόγου χάρι αν θέλει να βλέπει το βίντεο σε ολόκληρη την οθόνη ή όχι. Λόγω του ότι δεν επιτρέπεται η αυθαίρετη εναλλαγή αναπαραστάσεων, ο MPEG-DASH έχει Success Access Points (SAP) στα οποία μπορεί να αλλάζει η αναπαράσταση και μέσω των οποίων κάθε τμήμα ξεκινά. Μια αναπαράσταση μπορεί να περιέχει μόνο ένα από τα πιο κάτω:

- ένα ή περισσότερα segmentList,
- ένα segmentTemplate
- ένα ή περισσότερα BaseUrls, τουλάχιστον ένα SegmentBase, κανένα και κανένα segmentTemplate.

Πιο επεξηγηματικά, το segmentList περιέχει μια λίστα με segmentURLs, τα οποία πρέπει να αναπαραχθούν από τον client με την σειρά που εμφανίζονται στο MPD. Ένα στοιχείο segmentURL περιέχει μια διεύθυνση URL σε ένα τμήμα και πιθανώς ένα εύρος bytes. Μπορεί επίσης να υπάρχει ένα index segment στην αρχή του SegmentList. Το segmentBase δεν χρησιμοποιείται από τον MPEG-DASH, ενώ το segmentTemplate δημιουργεί μια λίστα από τμήματα δεδομένου ενός υπάρχοντος Template [46]. Η μορφή ενός MPD αρχείου καθορίζεται από ένα σύνολο περιορισμών που ονομάζονται profiles. Τα profiles μπορούν να γίνουν κατανοητά σαν άδεια για χρήστες του Dash και καθορίζονται από το ISO format [52]

4.2.3 Μορφή τμημάτων- Segments

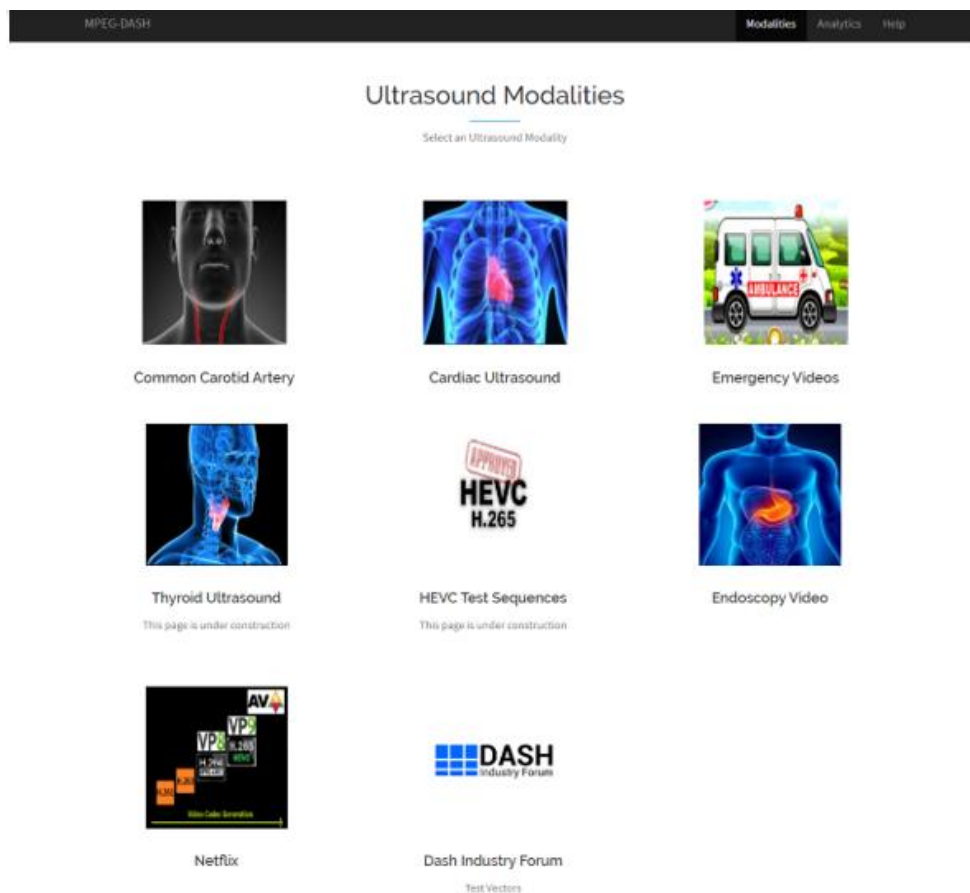
Τα τμήματα σε μια αναπαράσταση έχουν συνήθως ίδιο μήκος, παρότι το Mpeg Dash δεν περιορίζει το μήκος του τμήματος. Το περιεχόμενο των πολυμέσων μπορεί να υποδιαιρεθεί σε τμήματα. Ένα τμήμα μπορεί να θεωρηθεί η απάντηση ενός HTTP GET που κάνει ο Dash client. Ένα συστατικό από το πολυμέσο, για παράδειγμα το βίντεο, μπορεί να συμπιεστεί και να διαιρεθεί σε πολλά τμήματα [52]. Το πρώτο τμήμα μπορεί να είναι το τμήμα αρχικοποίησης, το οποίο περιέχει τις απαραίτητες πληροφορίες για την αρχικοποίηση του, ώστε να χρησιμεύσει στην αποκωδικοποίηση. Οπότε το πρώτο τμήμα δεν περιλαμβάνει πραγματικά δεδομένα από το βίντεο. Η ροή πολυμέσων διαιρείται σε ένα ή περισσότερα τμήματα, κάθε ένα από τα οποία έχει ένα μοναδικό URI, ένα δείκτη, χρόνο έναρξης και διάρκεια. Κάθε segment περιέχει τουλάχιστον ένα σημείο πρόσβασης το οποίο μπορεί να είναι τυχαίας προσπέλασης ή switch-to-point. Έτσι η αποκωδικοποίηση μπορεί να αρχίσει χρησιμοποιώντας δεδομένα μόνο από αυτό το σημείο, προς τα εμπρός. Η Μορφή η οποία έχουν τα τμήματα ορίζεται από το MPEG-DASH και βασίζεται στο ISO Base Media File Format. Για να ενεργοποιηθεί η λήψη των τμημάτων, τα τμήματα μπορούν να υποδιαιρεθούν σε υποτμήματα subsegments, τα οποία αντιπροσωπεύουν ένα μικρότερο σύνολο μονάδων πρόσβασης για το συγκεκριμένο τμήμα. Σε τέτοια περίπτωση υπάρχει ένα segment index για κάθε τμήμα που περιγράφει την θέση των υποτμημάτων. Ο Dash client μπορεί να χρησιμοποιήσει το index για να ζητήσει τα subsegments χρησιμοποιώντας HTTP GETS [47].



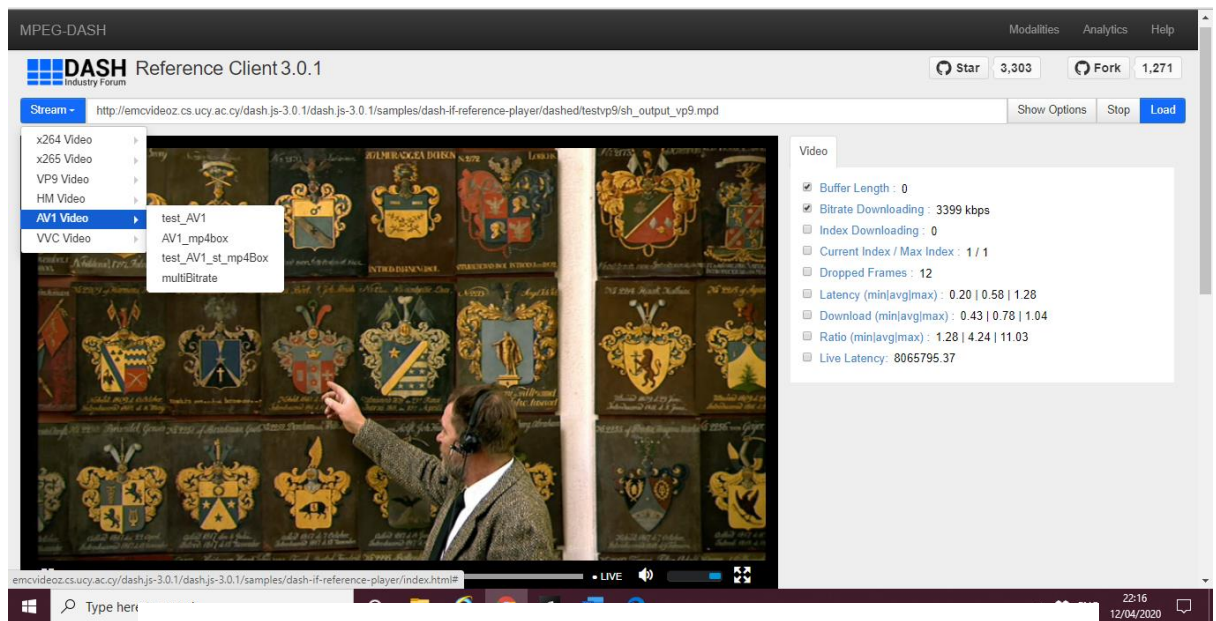
Σχήμα 4.2.3: Dash Data Model

4.3 Περιγραφή Συστήματος Προσαρμοστικής Μετάδοσης

Αφού έγινε η κατάλληλη επεξεργασία των βίντεο και κωδικοποίησή τους όπως αναφέρεται στο Κεφάλαιο 3, τα βίντεο αυτά ανέβηκαν στον Dash Player [55], όπου μπορούν να αναπαραχθούν και να αξιολογηθεί η ποιότητα τους. Αρχικά έγινε ανανέωση του Dash Player [47] στο τελευταίο version 3.0.1 από το παλιό 1.5.1 που είχε χρησιμοποιηθεί έως τώρα. Στο σχήμα 4.3.1 παρουσιάζεται η διεπιφάνεια μέσω της οποίας ο παραλήπτης έχει την δυνατότητα να επιλέξει κατηγορία βίντεο που επιθυμεί. Για παράδειγμα στην κατηγορία Netflix υπάρχουν τα βίντεο που χρησιμοποιήθηκαν για την πειραματική αξιολόγηση όπως περιγράφονται στην ενότητα 3.3. Στην κατηγορία Dash Industry Forum, βρίσκονται παραδείγματα λειτουργίας του MPEG-DASH, ενώ στις υπόλοιπες κατηγορίες παρουσιάζονται κυρίως υπερηχογραφήματα. Στην συνέχεια ο χρήστης μεταφέρεται στον Dash Player, στον οποίο έγιναν διάφορες τροποποιήσεις έτσι ώστε να εμφανίζονται τα βίντεο που ανήκουν στην κατηγορία που επιλέχθηκε, όπως παρουσιάζεται στο σχήμα 4.3.2.



Σχήμα 4.3.1: Διεπιφάνεια κατηγοριών



Σχήμα 4.3.2.: Διεπιφάνεια Dash Player

Dash Player

Το Dash player αναπτύχθηκε από το DASH Industry Forum, δημιουργώντας ένα ποιοτικό πλαίσιο, ποικίλων επιλογών έτσι ώστε να αναπαράγεται βίντεο του οποίου το περιεχόμενο είναι συμβατό με το MPEG-DASH. Ο Player λειτουργεί με Javascript βιβλιοθήκες που τρέχουν στην πλευρά του παραλήπτη. Τα βίντεο τα οποία παρουσιάζονται στον Player και ως εκ τούτου μπορούν να αναπαραχθούν, πέρασαν από τα στάδια της κωδικοποίησης, μέσω του ffmpeg, και της τμηματοποίησης, μέσω του MP4Box. Ουσιαστικά ο Player μπορεί να πάρει τις διευθύνσεις των MPD, μέσω ενός JSON αρχείου, έτσι ώστε να αναπαραχθούν τα βίντεο. Δημιουργήθηκε ένα JSON αρχείο για κάθε κατηγορία που μπορεί να επιλέξει ο χρήστης και η μορφή του παρουσιάζεται στο σχήμα 4.3.3. Σε ένα αρχείο JSON, ο αποστολέας έχει την δυνατότητα να προσθέσει την διεύθυνση url που βρίσκεται το βίντεο, και ένα όνομα (name) που του αντιστοιχεί. Ο παραλήπτης μπορεί να δει το url και το αντίστοιχο όνομα, στην λίστα με τα βίντεο τις διεπιφάνειας του Player, όπως παρουσιάζεται στο σχήμα 4.3.2.

```

{
  "items": [
    {
      "name": " x264 Video",
      "submenu": [
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx264/name25_dash.mpd", "name" : "test_x264", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx264/name_dash.mpd", "name" : "x264_50", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx264/sh_768x432_50_QP22_dash.mpd", "name" : "x264_sh", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx264/final2.mpd", "name" : "manifest", "browsers": "cdsbi" }
      ]
    },
    {
      "name": " x265 Video",
      "submenu": [
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx265/test_x265_onlyc_dash.mpd", "name" : "test_x265", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx265/output_x265_mc.mpd", "name" : "x265_mc_nocopy", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx265/output_x265.mpd", "name" : "test_x265_shaka", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx265/output.mpd", "name" : "test3", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testx265/output_dash.mpd", "name" : "test2", "browsers": "cdsbi" }
      ]
    },
    {
      "name": " VP9 Video",
      "submenu": [
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testvp9/sh_output_vp9.mpd", "name" : "VP9_sh_skaka_segment", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testvp9/exampleforvp9.mpd", "name" : "test_vp9_shaka", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testvp9/sh_vp9.mpd", "name" : "sh_vp9", "browsers": "cdsbi" }
      ]
    }
  ]
}

```

```

    },
    {
      "name": " HM Video",
      "submenu": [

        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testhm/tr_768x432_25_QP27_dash.mpd", "name" : "HM_mp4box", "browsers": "cdsbi" }

      ]
    },
    {
      "name": " AV1 Video",
      "submenu": [

        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testav1/test_copy_dash.mpd", "name" : "test_AV1", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testav1/tr_768x432_25_QP27_dash.mpd", "name" : "AV1_mp4box", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testav1/st_768x432_25_QP27_dash.mpd", "name" : "test_AV1_st_mp4Box", "browsers": "cdsbi" },
        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testav1/manifest/multi.mpd", "name" : "multiBitrate", "browsers": "cdsbi" }

      ]
    },
    {
      "name": " VVC Video",
      "submenu": [

        { "url" : "http://emcvideoz.cs.ucy.ac.cy/dash.js-3.0.1/dash.js-3.0.1/samples/dash-if-reference-player/dashed/testvvc/sh_768x432_50_QP22_dash.mpd", "name" : "test_VVC_sh_mp4Box", "browsers": "cdsbi" }

      ]
    }
  ]
}

```

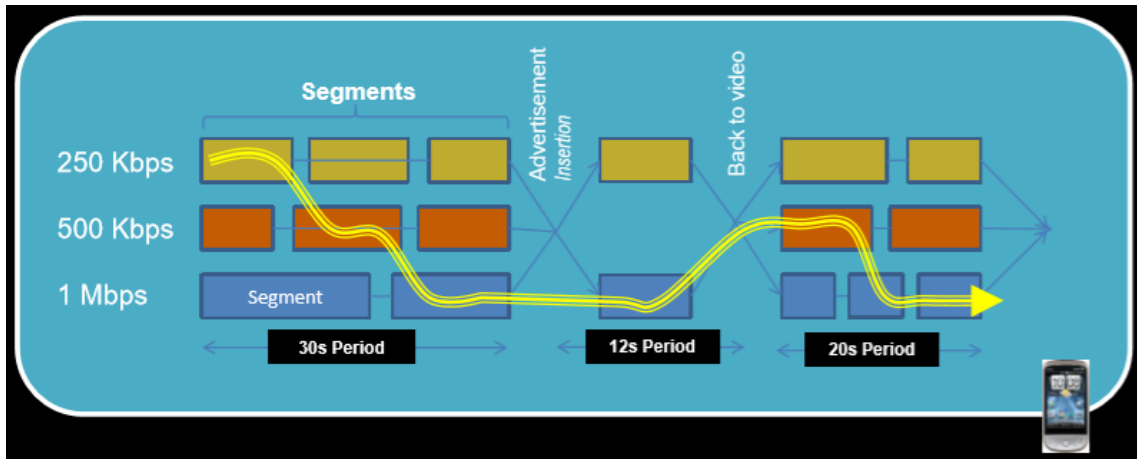
Σχήμα 4.3.3: Αρχείο JSON

Τα MPD αρχεία δημιουργήθηκαν με την βοήθεια του MP4Box και ελέγχθηκαν με την βοήθεια του Dash If Validator, όπως περιγράφεται στην ενότητα. Ουσιαστικά, η ιδιότητα του MP4Box να παρουσιάζει δεδομένα HTTP Adaptive Streaming, χρησιμοποιήθηκε στην παρούσα μελέτη. Αρχικά τα κωδικοποιημένα βίντεο μετατρέπονται σε .mp4 αρχεία μέσω του ffmpeg και το MP4Box για να δημιουργήσουμε τα τμήματα του βίντεο και το αντίστοιχο MPD αρχείο, με την εξής εντολή:

MP4Box -dash 4000 -frag 4000 -rap -segment-name segment_ output.mp4

Όπου κάθε τμήμα έχει διάρκεια 4000 ms και έχουμε μόνο ένα υποτμήμα ανά τμήμα. Το MPD αρχείο το οποίο θα πάρουμε σαν έξοδο είναι αυτό που θα δοθεί ως είσοδος στον Dash Player, μέσω του JSON αρχείου για να αναπαραχθεί.

Διαδικασία Προσαρμογής



Σχήμα 4.2.4: Η διαδικασία προσαρμογής που πραγματοποιείται στην πλευρά του παραλήπτη[56]

Η διαδικασία προσαρμογής πραγματοποιείται στο επίπεδο του παραλήπτη, ουσιαστικά ο παραλήπτης κατεβάζει το αρχείο MPD που επιλέγει, όπως στο σχήμα 4.2.2 [55] και κατεβάζονται διαδοχικά τμήματα του βίντεο βάση του playout buffer και του διαθέσιμου εύρους ζώνης το οποίο υπολογίζεται τοπικά. Στην συνέχεια επιλέγεται η κατάλληλη αναπαράσταση για να αναπαραχθεί. Οι παράγοντες που καθορίζουν ποια αναπαράσταση από το MPD αρχείο θα επιλεγεί καθορίζονται από την κατάσταση του buffer, τις συνθήκες του δικτύου, την επιλεγμένη ανάλυση του βίντεο (π.χ. αν θα αναπαραχθεί σε ολόκληρη οθόνη ή όχι) και από τους πόρους και άλλες δραστηριότητες της συσκευής [12].

Κεφάλαιο 5

Μέθοδοι Προσαρμοστικής Μετάδοσης Βίντεο

5.1	Μέθοδοι Προσαρμοστικής Μετάδοσης Βίντεο	66
5.2	Ασύρματες επικοινωνίες	67
	5.2.1 Προσομοίωση ασύρματων δικτύων	67
	5.2.2 Adaptive Controller και TruNET Simulator	68
5.3	Μέθοδος Προσαρμογής του Bitrate	70

Σε αυτό το κεφάλαιο θα επαληθεύσουμε τα αποτελέσματα για την σύγκριση κωδικοποιητών που έγιναν πιο πάνω και θα αναπαραστήσουμε πως γίνεται η προσαρμοστικότητα του bitrate των βίντεο, με την βοήθεια python scripts και του προσομοιωτή TruNET, καθορίζοντας μια προσέγγιση για ρεαλιστικές ασύρματες επικοινωνίες.

5.1 Μέθοδοι Προσαρμοστικής Μετάδοσης Βίντεο

Στις μέρες μας, το HTTP streaming έχει γίνει πλέον ένας αποδεκτός τρόπος για την μετάδοση πολυμέσων. Για να προσαρμοστεί ένα βίντεο στις απαιτήσεις του δικτύου, ο πάροχος δημιουργεί πολλές αναπαραστάσεις του αυθεντικού βίντεο. Υπάρχουν ποικίλοι μέθοδοι προσαρμογής που υποστηρίζουν adaptive HTTP streaming. Σε αυτό το κεφάλαιο διερευνούμε κάποιες τυπικές μεθόδους προσαρμογής και αξιολογούμε και συγκρίνουμε τα αποτελέσματα σε σχέση με την συμπεριφορά του bitrate. Στο τέλος, διαπιστώνουμε ότι η αντίληψη της προσαρμογής εξαρτάται όχι μόνο από την μέθοδο που ακολουθείται, αλλά και από το ίδιο το περιεχόμενο. Παρουσιάζεται επίσης η προετοιμασία των συνόλων αναπαραστάσης, η οποία μπορεί να επηρεάσει την συμπεριφορά κάποιων προσαρμοστικών μεθόδων [21].

Οι μέθοδοι προσαρμογής μπορούν να ταξινομηθούν από διαφορετικούς παράγοντες, όπως ο αριθμός των ροών και ο αριθμός των χρηστών. Κατά την μέθοδο προσαρμογής του bitrate, επιλέγεται δυναμικά η καλύτερη ποιότητα αναπαραγωγής βίντεο με βάση το διαθέσιμο εύρος ζώνης δικτύου και το μέγεθος του προγράμματος αναπαραγωγής βίντεο. Όταν το διαθέσιμο εύρος ζώνης είναι υψηλό, τότε το βίντεο μεταδίδεται σε υψηλότερη ποιότητα. Όταν το εύρος ζώνης μειωθεί, τότε θα μεταδοθεί μια έκδοση χαμηλότερης ποιότητας του βίντεο. Η ποιότητα και η ανάλυση του βίντεο εξαρτάται από πολλά χαρακτηριστικά, για παράδειγμα αν ένας χρήστης παρακολουθεί σε μικρή οθόνη, θα έχει πάντα μια μικρότερη έκδοση του βίντεο. Η προσαρμοστική ροή bitrate κάνει όλη αυτή την εργασία στο παρασκήνιο, ενώ το βίντεο αναπαραγεται με το λιγότερο buffering.

5.2 Ασύρματες επικοινωνίες

Με βάση μια σειρά τυπικών, βέλτιστων και χειρότερων σεναρίων, ο ασύρματος προσομοιωτής TruNET θα χρησιμοποιηθεί για τη μοντελοποίηση των χαρακτηριστικών σήματος σε όλες τις τοπολογίες ασύρματου δικτύου. Σε αυτή την εργασία στόχος είναι η δημιουργία ενός ρεαλιστικού περιβάλλοντος μοντελοποίησης για διερεύνηση των επιπτώσεων που προκύπτουν από διαφορετικά σενάρια πραγματικής ζωής. Στόχος είναι να καθοριστεί ο τρόπος διασύνδεσης και επικοινωνίας του προσαρμοστικού ελεγκτή (adaptive controller), και του ασύρματου προσομοιωτή TruNET. Σε αυτό το κεφάλαιο παρουσιάζεται μια ανάλυση της λειτουργικότητας του προσομοιωτή TruNET, καθώς και μια προτεινόμενη λύση διασύνδεσης.

5.2.1 Προσομοίωση Ασύρματων δικτύων

Οι ρεαλιστικές προσομοιώσεις δίνουν τη δυνατότητα στους ερευνητές να διερευνήσουν μια πληθώρα ασύρματων δικτύων που περιλαμβάνουν διαφορετικές ρυθμίσεις και τοπολογίες που δεν είναι εφικτές κατά τη χρήση της διαθέσιμης αναπτυγμένης υποδομής προμηθευτών κινητών τηλεπικοινωνιών ή / και κινητών ευρυζωνικών δικτύων.

Οι ερευνητές με την βοήθεια των ρεαλιστικών προσομοιώσεων μπορούν να διερευνήσουν πολλά ασύρματα δίκτυα τα οποία περιλαμβάνουν διαφορετικές ρυθμίσεις και τοπολογίες. Οι διαχειριστές δικτύου δεν μοιράζονται ακριβείς πληροφορίες σχετικά με το πως είναι διαμορφωμένο το δίκτυό τους, ενώ οι τοπολογίες εδάφους θέτουν

σημαντικές προκλήσεις όσο αφορά την επικύρωση των προτεινόμενων προσαρμοστικών αλγορίθμων μετάδοσης βίντεο [57].

Ο TruNET είναι ένας ασύρματος προσομοιωτής βραβευμένος για τις προσομοιώσεις δικτύων κινητής ευρυζωνικής επικοινωνίας. Διερευνά στιγμιαία τη διαθεσιμότητα του εύρους ζώνης προσομοιώνοντας την διάδοση σήματος και την εξασθένησή του, και δίνει μετρήσεις QoS οι οποίες αφορούν την διαμόρφωση του δικτύου. Ο στόχος είναι να εξαχθούν στατιστικά στοιχεία μεταδεδομένων, τα οποία θα βοηθήσουν στην βελτίωση του προσαρμοστικού μηχανισμού ελέγχου του βίντεο, έτσι ώστε να αποφευχθεί το φαινόμενο ping pong [58].

5.2.2 Adaptive Controller και TruNET Simulator

Ο προσομοιωτής TruNET, μπορεί να προσομοιώσει τα χαρακτηριστικά μιας ραδιοφωνικής μετάδοσης, οποιουδήποτε ασύρματος συστήματος [57], και την αξιολόγηση της απόδοσης WiFi σε δίκτυα 802.11. Η παροχή ενός συγκεκριμένου επιπέδου κάλυψης σε ένα συγκεκριμένο σύστημα είναι βάση για την διατήρηση μιας σταθερού επιπέδου υπηρεσίας που οδηγεί σε υψηλές ή χαμηλές ταχύτητες ασύρματης σύνδεσης. Ανάλογα με τους ρυθμούς απόδοσης της ασύρματης σύνδεσης μπορεί να βελτιωθεί η απόδοση της μετάδοσης βίντεο [58].

Το ασύρματο TruNET, μπορεί να παράγει γεωγραφικούς χάρτες απόδοσης, οι οποίοι αναφέρονται στην ρεαλιστική απόδοση των τυπικών ασύρματων συστημάτων, σε ένα σενάριο ενδιαφέροντος. Στην παρούσα μελέτη θα εξεταστούν δυο σενάρια, ένα που αφορά την χρήση δικτύου από πολλούς χρήστες και ένα που αφορά την χρήση δικτύου από μόνο ένα χρήστη. Τα σχήματα παρακάτω παρουσιάζει τα αποτελέσματα προσομοίωσης τέτοιων εσωτερικού σεναρίων.



Σχήμα 5.2.1: Χάρτης δικτύου ενός χρήστη



Σχήμα 5.2.2: Χάρτης δικτύου πολλαπλών χρηστών

Με βάση την λειτουργικότητα του TruNET, που αναφέρθηκε στα προηγούμενα κεφάλαια, το TruNET θα παράγει χάρτες που αναφέρονται στο Throughput, ανάλογα με το σενάριο, όπως αυτά που παρουσιάζονται στο σχήμα. Στη συνέχεια, το TruNET, θα εξάγει αυτούς τους χάρτες σε μορφή CSV, για να μπορούν να χρησιμοποιηθούν από την μονάδα προσαρμοστικού ελέγχου.

Η προσαρμοστική μονάδα ελέγχου θα εφαρμοστεί και θα επικυρωθεί στη γλώσσα Python, χρησιμοποιώντας ακριβείς πληροφορίες χρονισμού που είναι διαθέσιμες τόσο στα βίντεο που δημιουργήθηκαν, όσο και στους χάρτες TruNET. Για την υλοποίηση της προσαρμοστικής απόφασης ελέγχου (adaptive control decision) θα υποστηρίζονται τρεις τρόποι λειτουργίας. Όλες οι προσεγγίσεις θα βασίζονται στην τεχνική “table-lookup and match”, όπου οι γεωχωρικοί χάρτες απόδοσης που δημιουργούνται από το ασύρματο TruNET, θα παρέχουν χρονικές μεταβολές στιγμιαίας απόδοσης. Ο προσαρμοστικός ελεγκτής θα αντιδρά σε αυτές τις πληροφορίες για την παροχή βελτιστοποιημένων υπηρεσιών μετάδοσης βίντεο, που μπορούν να υποστηρίζονται ρεαλιστικά μέσω συγκεκριμένων ασύρματων καναλιών ή συνδέσμων που αναφέρονται σε συγκεκριμένες γεωχωρικές διαδρομές [58].

5.3 Μέθοδος Προσαρμογής του Bitrate

Σε αυτή την μελέτη, εστιάζουμε στις μεθόδους προσαρμογής για την περίπτωση ενός και πολλαπλών χρηστών, που βασίζονται στην απόδοση του δικτύου (Throughput). Για να γίνει αυτό πάρθηκαν μετρήσεις της απόδοσης του δικτύου (Throughput), σε Mbps με την βοήθεια του TruNET Network simulation. Έχουμε τιμές για ένα χρήστη, όπως φαίνεται στο σχήμα 5.2.1 αλλά και για πολλούς χρήστες, στο σχήμα 5.2.2 όταν χρησιμοποιούν το δίκτυο, για να μπορούμε να μελετήσουμε και τις δύο περιπτώσεις.

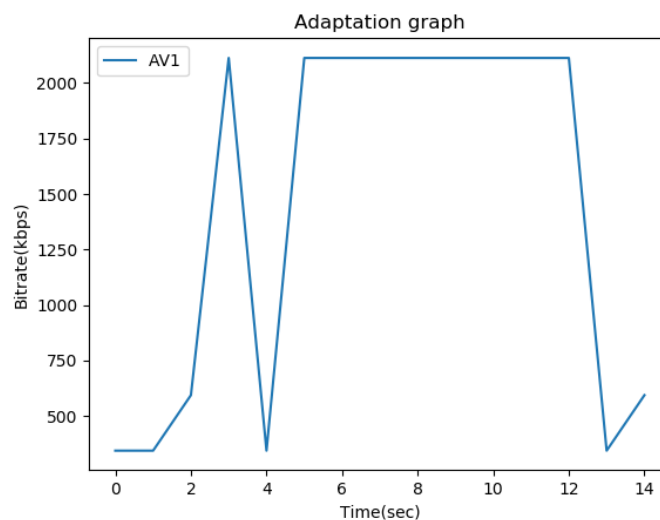
Οι τιμές είναι καταγεγραμμένες σε ένα Excel αρχείο, το οποίο διαβάζεται μέσω ενός python Script, και συγκρίνεται με τα διαθέσιμα bitrate των κωδικοποιημένων βίντεο, ώστε να ελέγξουμε ποιο από αυτά τα bitrate, ταιριάζει καλύτερα στην απόδοση του δικτύου. Επειδή το εύρος ζώνης σύνδεσης είναι πολύ κυμαινόμενο, ο ρυθμός bitrate βίντεο, άρα και η ποιότητά του, πρέπει να ρυθμιστεί γρήγορα. Οι αλλαγές του ρυθμού bitrate επηρεάζουν την αντίληψη των τελικών χρηστών. Υλοποιήθηκαν τρεις διαφορετικές περιπτώσεις προσαρμογής, οι οποίες επεξηγούνται πιο κάτω:

Αρχικά, τρέχοντας το αρχείο AdaptationScripts.py ο χρήστης έχει την δυνατότητα να επιλέξει ποιο αλγόριθμο προσαρμογής θα εκτελεστεί και στην συνέχεια αν η προσαρμογή θέλει να γίνει για ένα ή για πολλούς χρήστες, καθώς και το πρότυπο κωδικοποίησης που θα μελετηθεί. Οι κώδικες όλων των αρχείων python που δημιουργήθηκαν παρουσιάζονται στο παράρτημα.

Το πρώτο python script που δημιουργήθηκε, συγκρίνει το διαθέσιμο Throughput, με το bitrate των βίντεο που χρησιμοποιήθηκαν κατά την πειραματική αξιολόγηση, διαβάζοντας από 2 αρχεία Excel. Η βασική ιδέα για την υλοποίηση αυτού του script είναι ότι επιλέγεται η κωδικοποίηση βίντεο η οποία έχει το μεγαλύτερο PSNR, το οποίο ταιριάζει στο διαθέσιμο bitrate. Στο σχήμα 5.2.3 παρουσιάζεται ένα παράδειγμα όπου φαίνεται η προσαρμογή του bitrate, και της αντίστοιχης τιμής του PSNR ανάλογα με το διαθέσιμο εύρος ζώνης. Παρατηρείται ότι όταν το διαθέσιμο Throughput είναι 1100 kbps, επιλέγεται ένα πιο μικρό bitrate, 685.8 kbps, ενώ όταν το Throughput αυξάνεται, επιλέγεται ένα πιο μεγάλο bitrate, αλλά όχι μεγαλύτερο από αυτό του Throughput. Επίσης, στο σχήμα 5.2.4 παρουσιάζεται μια γραφική παράσταση, που δείχνει τις αλλαγές στο bitrate, σε σχέση με τον χρόνο, για αυτή την κλασική μορφή προσαρμογής.

```
Current Time: 1
Available Throughput: 1100.63
Selected bitrate: 685.759 kbps
Selected psnr: 38.296 dB
Current Time: 2
Available Throughput: 3400.95
Selected bitrate: 1579.515 kbps
Selected psnr: 41.305 dB
Current Time: 3
Available Throughput: 3300.12
Selected bitrate: 1579.515 kbps
Selected psnr: 41.305 dB
Current Time: 4
```

Σχήμα 5.2.3 Bitrate Adaptation

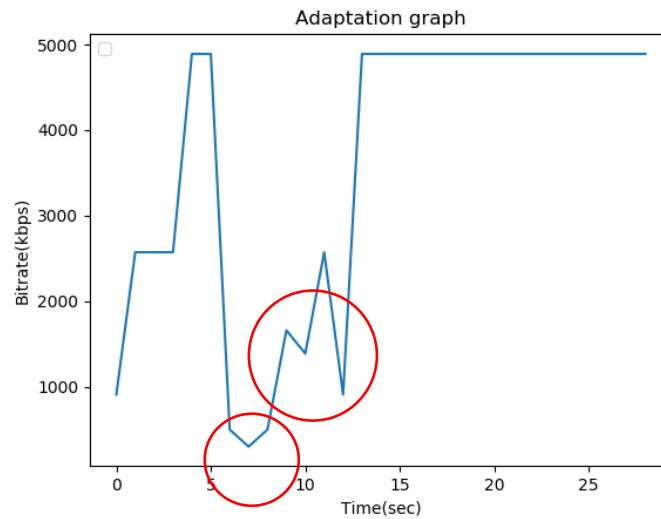


Σχήμα 5.2.4: Classic Adaptation for many users, Av1 codec.

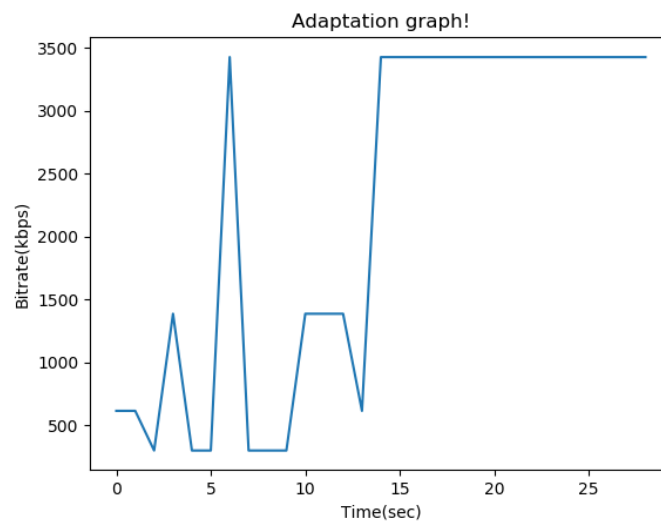
Στο δεύτερο python Script, τα δεδομένα όσο αφορά την κωδικοποίηση του αντί να διαβάζονται από ένα excel αρχείο, διαβάζονται από ένα αρχείο MPD. Από το αρχείο διαβάζεται το bitrate κάθε αναπαράστασης, καθώς και το ID της, για να ξέρουμε σε ποια κωδικοποίηση αναφερόμαστε. Συγκρίνει το available Throughput, με το bitrate των βίντεο που έχουμε και επιλέγεται το ID της κωδικοποίησης που ταιριάζει στο διαθέσιμο bitrate, η διαδικασία που ακολουθείται εδώ ταυτίζεται με την διαδικασία υλοποίησης στο πρώτο script.

Στο τρίτο python Script, παρουσιάζεται μια διαφορετική προσέγγιση της κλασσικής προσαρμοστικής μετάδοσης, ουσιαστικά με αυτό το τρόπο θέλουμε να αποφύγουμε το ping pong effect πιθανόν να δημιουργείται, δηλαδή οι επαναλαμβανόμενες εναλλαγές του bitrate από ψηλότερη σε χαμηλότερη τιμή και το ανάποδο. Την πρώτη φορά βρίσκουμε την κωδικοποίηση που ταιριάζει το διαθέσιμο Throughput. Στην συνέχεια κάθε φορά που το διαθέσιμο Throughput αλλάζει, συγκρίνεται με την προηγούμενη διαθέσιμη τιμή. Αν η νέα τιμή είναι μικρότερη από 20%, δεν γίνεται έλεγχος για να αλλάξει το bitrate της κωδικοποίησης. Στην περίπτωση που η νέα τιμή του Throughput είναι μεγαλύτερη από 20%, τότε αφαιρούμε το 50% της τιμής του Throughput και μετά επιλέγεται το κατάλληλο bitrate, έτσι αποφεύγονται скаμπανεβάσματα στις τιμές του bitrate, το ping pong effect που πιθανόν να δημιουργείται, στην πρώτη λύση και δεν υπάρχει μεγάλο overhead. Στην περίπτωση που όλα τα bitrate είναι μεγαλύτερα από το available Throughput, τότε επιλέγεται το πιο μικρό bitrate και το αντίστοιχο PSNR.

Για την επαλήθευση αυτής της μεθόδου, χρησιμοποιήθηκαν στοχευμένες τιμές για το Throughput και ποιο κάτω παρουσιάζεται ένα παράδειγμα της κωδικοποίησης x265, όπου με την αρχική υλοποίηση υπάρχει το ping pong effect στο σχήμα 5.2.5, ενώ όταν τρέχουμε το δεύτερο script φαίνεται ότι υπάρχει πιο ομαλή προσαρμογή του bitrate, στο σχήμα 5.2.6

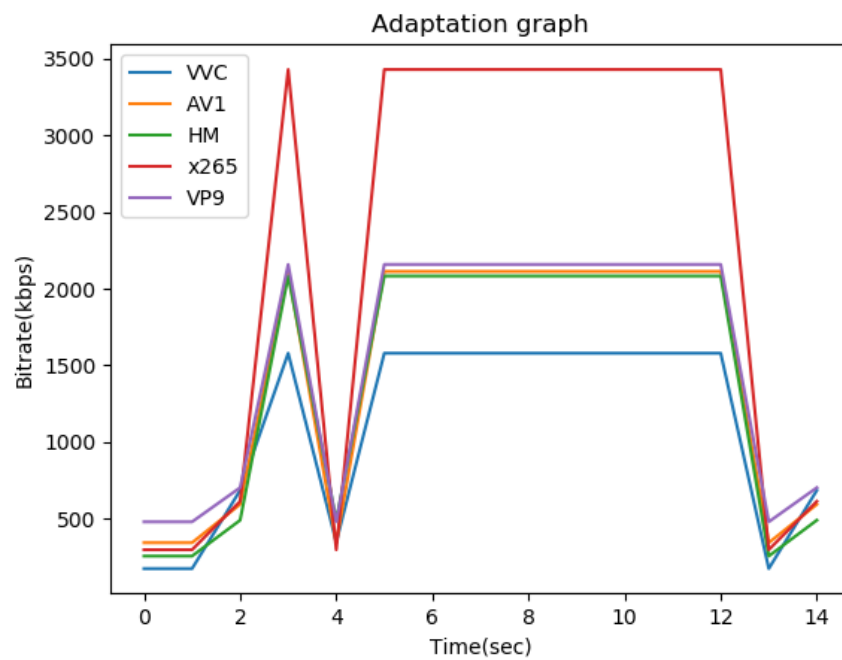


Σχήμα 5.2.5: x265 Adaptation-one user - Ping Pong effect



Σχήμα 5.2.6: x265 Adaptation-one user – Solve Ping Pong effect

Στην γραφική παράσταση 5.2.7 παρουσιάζεται η προσομοίωση της προσαρμογή του bitrate ενός βίντεο σε σχέση με τον χρόνο για όλους τους κωδικοποιητές. Παρατηρούμε ότι ο VVC, δίνει τα καλύτερα αποτελέσματα σε bitrate, και το x265 τα χειρότερα αποτελέσματα. Επίσης τα αποτελέσματα του AV1 και του HM είναι πάρα πολύ κοντά, και ακολουθεί το VP9. Τα παραπάνω επαληθεύουν τα αποτελέσματα του εξάχθηκαν στο κεφάλαιο 3.



Σχήμα 5.2.7: All Codecs, many users, classic adaptation

Κεφάλαιο 6

Συμπεράσματα

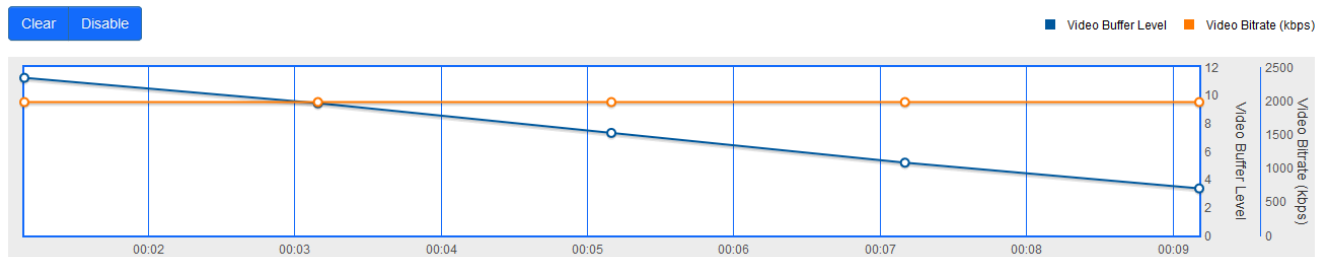
6.1	Συμπεράσματα για το Σύστημα Προσαρμοστικής Μετάδοσης	75
6.2	Συμπεράσματα για τα Αποτελέσματα από την Σύγκριση Κωδικοποιητών	77
6.3	Μελλοντική Εργασία	78

Σε αυτό το κεφάλαιο παρατίθενται συνοπτικά τα κυριότερα συμπεράσματα που εξάχθηκαν από τα αποτελέσματα, όσο αφορά της σύγκριση των προτύπων κωδικοποίησης καθώς και το σύστημα προσαρμοστικής μετάδοσης, και την δοκιμή του με την βοήθεια των αλγορίθμων προσαρμογής. Τέλος, παρουσιάζεται η μελλοντική εργασία που μπορεί να γίνει.

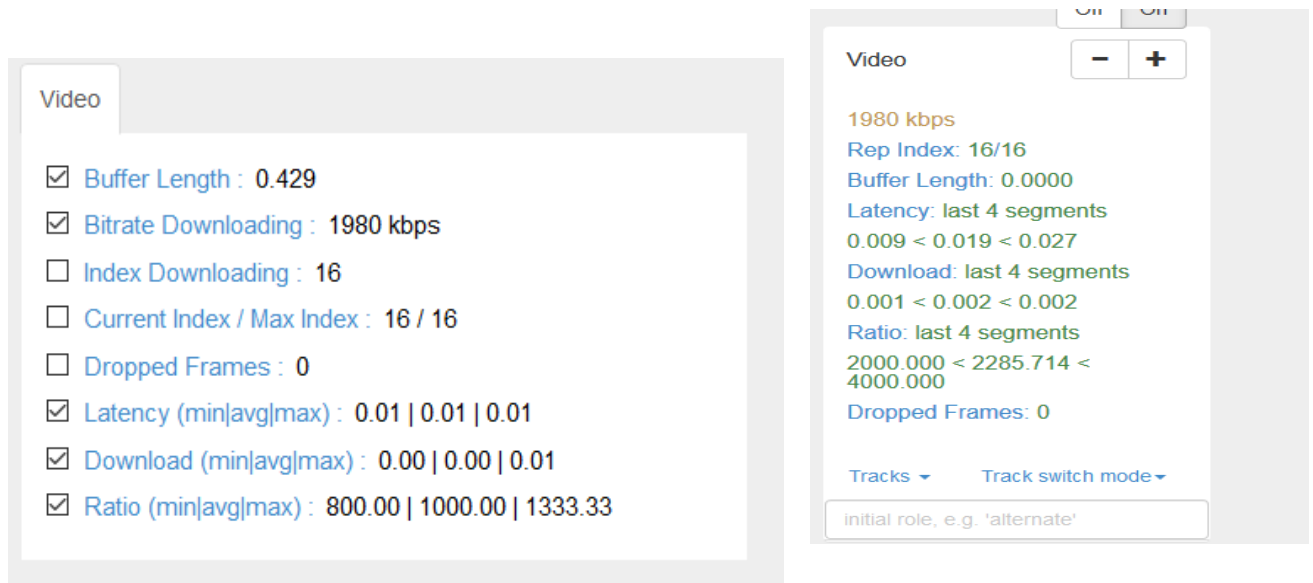
6.1 Συμπεράσματα για το Σύστημα Προσαρμοστικής Μετάδοσης

Το σύστημα προσαρμοστικής μετάδοσης βίντεο με την βοήθεια του Dash Player 3.0.1 του Dash Industry Forum, παρουσιάζει μεγάλες αλλαγές σε σχέση με την προηγούμενη έκδοση που είχε χρησιμοποιηθεί του Dash Player Reference Client 1.5.1. Το προηγούμενο σύστημα αφορούσε την ιατρική εκπαίδευση και αξιολόγηση ιατρικού βίντεο. Το παρών σύστημα συνεχίζει να χρησιμοποιείται για σκοπούς εκπαίδευσης ανθρώπων που ενδιαφέρονται να μάθουν για υπερηχογραφήματα αλλά και για αξιολόγηση βίντεο από το σύνολο δεδομένων του Netflix, καθώς επίσης έχουν και την ευκαιρία να εξετάσουν περισσότερες πληροφορίες για το βίντεο κατά την διάρκεια του streaming, όπως για παράδειγμα γραφικές παραστάσεις που αφορούν το Video Buffer Level, Bitrate (kbps), Ratio, Download Rate(Mbps), Latency(ms), ένα παράδειγμα παρουσιάζεται στο σχήμα 6.1. Η ανανεωμένη έκδοση του Dash Player 3.0.1 δίνει καλύτερα αποτελέσματα κωδικοποίησης και τα βίντεο έχουν καλύτερη ποιότητα από ότι η παλαιότερη έκδοση, όπως για παράδειγμα καλύτερο Video Download Rate και Video

Ratio, όπως συγκρίνονται τα αποτελέσματα για το ίδιο βίντεο στα σχήματα 6.2 και 6.3. Ο Dash Player 1.5.1 λειτουργούσε μόνο σε Internet Explorer, και χρησιμοποιήθηκε μόνο το πρότυπο κωδικοποίησης x264, ενώ ο 3.0.1 λειτουργεί σε όλα τα προγράμματα περιήγησης. Όσο αφορά το σύνολο βίντεο από το Netflix, ο Dash Player μπορεί να αναπαράξει βίντεο που είναι κωδικοποιημένα με τα πρότυπα x264, VP9 και AV1 σε όλα τα προγράμματα περιήγησης, εκτός από το AV1 που δεν δουλεύει στο Google Chrome, ενώ δεν μπορεί να αναπαράξει τα βίντεο με κωδικοποίηση HM και VVC.



Σχήμα 6.1: Γραφική παράσταση Video Buffer Level και του Video Bitrate (kbps) από τον Dash Player 3.0.1



Σχήμα 6.2: Παρουσιάζονται πληροφορίες όσο αφορά τις παραμέτρους προσαρμοστικής μετάδοσης, αριστερά για τον Video Streaming Dash Player 3.0.1 και δεξιά για το Video Streaming Dash Player 1.5.1

Τα συμπεράσματα που μπορούμε να εξάγουμε όσο αφορά τις μεθόδους προσαρμοστικής μετάδοσης είναι ότι μέσα από τους αλγορίθμους προσαρμογής που έγιναν σε γλώσσα python, επιτεύχθηκε να επιλέγεται δυναμικά η καλύτερη ποιότητα αναπαραγωγής βίντεο με βάση το διαθέσιμο εύρος ζώνης δικτύου. Όταν το διαθέσιμο εύρος ζώνης είναι υψηλό, τότε το βίντεο μεταδίδεται σε υψηλότερη ποιότητα. Όταν το εύρος ζώνης μειωθεί, τότε θα μεταδοθεί μια έκδοση χαμηλότερης ποιότητας του βίντεο. Επίσης χρησιμοποιώντας ένα Threshold στο Throughput, μπορεί να αποφυγή το ping pong effect.

6.2 Συμπεράσματα για τα Αποτελέσματα από την Σύγκριση Κωδικοποιητών

Συμπερασματικά, από την μελέτη των πινάκων τιμών του BD-Rate για το PSNR, τις συγκρίσεις με άλλες αναφορές και από την προσομοίωση της προσαρμοστικής μετάδοσης παρατηρούμε ότι το πρότυπο κωδικοποίησης VVC, ξεπερνά σταθερά τους κωδικοποιητές HM,x265,VP9 και HM και δίνει σημαντικά κέρδη και απόδοση κωδικοποίησης. Ωστόσο, το VVC, είναι ακόμα στα αρχικά του στάδια και απαιτεί την χρήση περισσότερων εργαλείων, για να οριστικοποιηθεί και να βελτιστοποιηθεί πλήρως έτσι ώστε να ανταγωνιστεί αποτελεσματικά το AV1 της libaom, που είναι πιθανώς το μέλλον της βιομηχανίας του video streaming.

Τα συμπεράσματα αυτά μπορούν να επικυρωθούν και μέσω του προσαρμοστικού μηχανισμού ελέγχου, όπου προσημειώνεται η διαδικασία προσαρμογής του bitrate, σε διαθέσιμες τιμές του εύρους ζώνης, τις οποίες βρίσκουμε από τον ασύρματο προσομοιωτή TruNET.

Τέλος, Μέσα από την βιβλιογραφία, μελέτες που αφορούν βίντεο γενικής χρήσης δείχνουν ότι η απόδοση κωδικοποίησης επηρεάζεται σε μεγάλο βαθμό από το περιεχόμενο και το μέγεθος του βίντεο, αυτό αποδεικνύεται και από την συγκεκριμένη μελέτη. Για παράδειγμα, αυτό μπορούμε να το διαπιστώσουμε παρατηρώντας ότι σε μερικά βίντεο έχουμε αρκετά μεγάλες μετρήσεις Bitrate (kbps) , σε σχέση με τα υπόλοιπα βίντεο, γιατί στα συγκεκριμένα βίντεο το περιεχόμενό ασφαλή τους έχει αυξημένη κινητικότητα.

Επομένως ικανοποιητικά συμπεράσματα μπορούν να εξαχθούν μόνο όταν διεξαχθούν πειράματα που εκτείνονται σε όλη την κλίμακα ανάλυσης βίντεο και διαφορετικών περιεχομένων. Για παράδειγμα, εάν θέλουμε να επικεντρωθούμε στην ανάλυση ιατρικών βίντεο θα χρησιμοποιήσουμε ένα επαρκή αριθμό βίντεο ιατρικού περιεχομένου.

6.3 Μελλοντική Εργασία

Όπως παρατηρήθηκε από τα αποτελέσματα, η κωδικοποίηση επηρεάζεται σε μεγάλο βαθμό από το περιεχόμενο του βίντεο. Βίντεο στα οποία υπήρχε μεγάλη κίνηση, παρουσίαζαν μεγάλο bitrate. Οπότεν είναι σημαντικό να χρησιμοποιηθεί ένα μεγαλύτερο δείγμα από βίντεο για την σύγκριση των προτύπων κωδικοποίησης, όπου θα υπάρχουν βίντεο με διαφορετικό περιεχόμενο, περισσότερες αναλύσεις και framerate.

Επίσης, στο μέλλον μπορούν να εφαρμοστούν περισσότερες τεχνικές αξιολόγησης ποιότητας βίντεο, όπως το SSIM, το VMAF, το οποίο χρησιμοποιείται από το Netflix, και χρησιμοποιεί ένα συνδυασμό μηχανικής μάθησης και πολλών μετρικών όπως η οπτική ποιότητα και η οπτική ποιότητα, καθώς και blind video quality assessment metrics (no-reference to original video). Η τελευταία μετρική είναι χρήσιμη όταν δεν υπάρχει πρόσβαση στο αρχικό βίντεο, καθώς μόνο από το βίντεο που λαμβάνεται από τον παραλήπτη μπορεί να γίνει αξιολόγηση και να υπολογιστεί η ποιότητα του, δεν μπορεί να γίνει σύγκριση με το αρχικό βίντεο.

Στην συνέχεια, σκοπός είναι να δημιουργηθεί ένα συνολικό σύστημα, με όλες τις Μεθόδους Προσαρμογής, όπου για την κωδικοποίηση βίντεο θα χρησιμοποιούνται περισσότεροι παραμέτροι QP, ώστε να υπάρχει ένα πιο ευρύ φάσμα σύγκρισης.

Τέλος, εκτός από το MPEG-DASH, όπου χρησιμοποιούνται προ-κωδικοποιημένα βίντεο, μπορεί να υλοποιηθεί το Adaptive Encoding, δηλαδή ανάλογα με το bitrate, μέσω ενός αλγορίθμου να χρησιμοποιείται η κατάλληλη κωδικοποίηση, που ταιριάζει στο διαθέσιμο bitrate.

Βιβλιογραφία

- [1] A. Panayides, M. S. Pattichis, C. S. Pattichis, C. P. Loizou, M. Pantziaris, and A. Pitsillides, “Atherosclerotic plaque ultrasound video encoding, wireless transmission, and quality assessment using h.264,” IEEE Transactions on Information Technology in Biomedicine, vol. 15, no. 3, pp. 387–397, 2011.
- [2] Z. C. Antoniou, A. S. Panayides, M. Pantzaris, A. G. Constantinides, C. S. Pattichis, and M. S. Pattichis, “Real-time adaptation to time-varying constraints for medical video communications,” IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, pp. 1177–1188, 2018.
- [3] A. Panayides, M. Pattichis, C. Loizou, M. Pantziaris, A. Constantinides, and C. Pattichis, “An effective ultrasound video communication system using despeckle filtering and hevc,” IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 2, pp. 668–676, March 2015.
- [4] Cisco, V.N.I. “Cisco Annual Internet Report (2018–2023)” , Networking Solutions White Paper, Cisco, San Jose, CA, USA, March, 2020.
- [5] Verge. “Disney plus and facebook are also reducing streaming quality in europe.”, 2020 Available: <https://www.theverge.com/2020/3/22/21189920/disney-plus-delayedfrance-uk-streaming-reduce-quality-coronavirus>
- [6] YouTube. “Youtube recommended upload encoding settings.”, 2016 Available: [https:// support.google.com/youtube/answer/1722171?hl=en](https://support.google.com/youtube/answer/1722171?hl=en).
- [7] Aristidou Andreas EPL 426: Computer Graphics, Cyprus: University of Cyprus, Chapter 17: Animation, 2020 Available: http://www.cs.ucy.ac.cy/courses/EPL426/courses/Lectures/17_Character_Animation.pdf
- [8] Robert M. Steinman, Zygmunt Pizlo, Filip J. Pizlo, “Phi is not beta, and why Wertheimer’s discovery launched the Gestalt revolution”, Vision Research, Vol. 40, Issue 17, 2000.
- [9] Λαζαρίνης, Φ. Τεχνολογίες Πολυμέσων: Θεωρία Ψηφιακού Βίντεο, Υλικό, Λογισμικό. Αθήνα: Εκδόσεις Κλειδάριθμος, 2007.
- [10] “Adaptive Streaming”, Video Encoding, Bitmovin, Available: <https://bitmovin.com/adaptive-streaming/>

- [11] Phillip A. Laplante, Multimedia Image Second Edition, Pennsylvania: Taylor & Francis Group, LLC, 2012.
- [12] Παναγίδης Αντρέας, ΠΕΣ 612: ΚΙΝΗΤΗ ΚΑΙ ΔΙΑΧΥΤΗ ΥΠΟΛΟΓΙΣΤΙΚΗ, Cyprus: Open University of Cyprus, Εφαρμογές Βίντεο στα πλαίσια της ΚΔΥ, 2017.
- [13] Νεοφύτου Μάριος, ΕΠΛ 445 Ψηφιακή Επεξεργασία Εικόνας, Cyprus: University of Cyprus, Κεφάλαιο 10 Ψηφιακό Video, 2019.
- [14] Ζεϊναλιπούρ Δημήτρης, ΕΠΛ 232: Προγραμματιστικές τεχνικές και Εργαλεία, Διάλεξη 17: Ανάπτυξη Λογισμικού σε Ομάδες και Λογισμικό Ανοικτού Πηγαίου Κώδικα Cyprus: University of Cyprus, 2018. Available: <https://www.cs.ucy.ac.cy/~dzeina/courses/epl232/lectures/17.pdf>
- [15] “FFmpeg, A complete, cross-platform solution to record, convert and stream audio and video”, Last Update: April 2020 Available: <https://www.ffmpeg.org/>
- [16] “FFmpeg, A complete, cross-platform solution to record, convert and stream audio and video”, wiki: ffmpeg Documentation, Last Update: March 2019, Available: <https://trac.ffmpeg.org/wiki/ffmpeg>
- [17] Dash IF Validator, Dash Industry Forum, Available: <https://conformance.dashif.org/>
- [18] A. S. Panayides, M. S. Pattichis, M. Pantziaris, A. G. Constantinides and C. S. Pattichis, "The Battle of the Video Codecs in the Healthcare Domain - A Comparative Performance Evaluation Study Leveraging VVC and AV1," IEEE Access, vol. 8, 2020.
- [19] Alliance for Open Media. Accessed: Sep. 2019. [Online]. Available: <https://aomedia.org/>
- [20] “Video Developer Report 2018”, Bitmovin, 2018. Available: <https://go.bitmovin.com/hubfs/Bitmovin-Video-Developer-Report-2018.pdf?submissionGuid=f41efe3c-7a1b-465e-8a1e-8eab839e4ad1>
- [21] Gangadharan Esakki, “ Adaptive Encoding for Constrained Video Delivery in HEVC, VP9, AV1 and VVC Compression Standards and Adaptation to Video Content”, The University of New Mexico, July 2020
- [22] “x265”, Wikipedia: The free Encyclopedia, 16th May 2020. Available: <https://en.wikipedia.org/wiki/X265>

- [23] Google. (2013) Vp9 video codec. Available: <https://storage.googleapis.com/downloads.webmproject.org/docs/vp9/vp9-bitstream-specification-v0.620160331-draft.pdf>
- [24] J. Bankoski, P. Wilkins, and Y. Xu, “Technical overview of vp8, an open source video codec for the web,” in 2011 IEEE International Conference on Multimedia and Expo. IEEE, 2011, pp. 1–6.
- [25] C. G. Bampis, Z. Li, and A. C. Bovik, “Spatiotemporal feature integration and model fusion for full reference video quality assessment,” IEEE Transactions on Circuits and Systems for Video Technology, 2018.
- [26] MacDougall Paul, “Get Ready for a Multi-Codec World”, Bitmovin, 12th July, 2018. Available: <https://bitmovin.com/get-ready-for-a-multi-codec-world/>
- [27] Jankkoi, “How netflix delivers better-looking downloads without eating up all your phone storage.”, 2019. Available: <https://variety.com/2016/digital/news/netflix-offlinedownloads-codecs-vp9-1201932502/>.
- [28] Versatile Video Coding ,VVC, 2018, Fraunhofer Heinrich Institute, Available: <https://jvet.hhi.fraunhofer.de/>
- [29] HEVC Test Model, HM, Synopsis, Available: <https://www.openhubs.net/p/hevc>
- [30] High Efficiency Video Coding ,HEVC, Fraunhofer Heinrich Institute, Available: <https://mpeg.chiariglione.org/tags/hm>
- [31] Collins Kayla “AV1 vs HEVC: Should You Abandon HEVC Now?”,Digiarty VideoProc, 13th March 2020, Available: <https://www.videoproc.com/media-converter/av1-vs-hevc.htm>
- [32] D. Grois, D. Marpe, A. Mulayoff, B. Itzhaky, and O. Hadar, “Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders,” in Proc. Picture Coding Symp. (PCS), San Jose, CA, USA, Dec. 2013, pp. 394–397.
- [33] D. Grois, T. Nguyen, and D. Marpe, “Coding efficiency comparison of AV1/VP9, H.265/MPEG-HEVC, and H.264/MPEG-AVC encoders,” in Proc. Picture Coding Symp. (PCS), 2016, pp. 1–5.

- [34] D. Grois, T. Nguyen, and D. Marpe, "Performance comparison of AV1, JEM, VP9, and HEVC encoders," *Proc. SPIE Appl. Digit. Image Process. XL*, vol. 10396, Feb. 2018, Art. no. 103960L.
- [35] P. Akyazi and T. Ebrahimi, "Comparison of compression efficiency between HEVC/H.265, VP9 and AV1 based on subjective quality assessments," in *Proc. 10th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Cagliari, Italy, May 2018, pp. 1–6.
- [36] J. L. Tanou and M. Blestel, "Analysis of emerging video codecs: Coding tools, compression efficiency," *SMPTE Motion Imag. J.*, vol. 128, no. 10, pp. 14–24, Nov. 2019.
- [37] "Improvements of the BD-PSNR model", Video Coding Experts Group (VCEG), ITU - Telecommunications Standardization Sector, July, 2008
- [38] G. Dan, N. Tung, and M. Detlev, "Coding efficiency comparison of av1/vp9, h. 265/mpeg-hevc, and h. 264/mpeg-avc encoders," in *2016 Picture Coding Symposium (PCS)*. IEEE, 2016, pp. 1–5.
- [39] C. T. Anatoliy Zabrovskiy, Christian Feldmann, "Multi-codec dash dataset," in *2015 IEEE International Conference on Image Processing (ICIP)*. ACM MMSys18, 2018.
- [40] "Performance comparison of av1, and jem, and vp9, and hevc encoders," in *Applications of Digital Image Processing XL*, vol. 10396. SPIE:International Society for Optics and Photonics, 2018, p. 103960L
- [41] Y. Liu. (2018) Av1 beats x264 and libvpx-vp9 in practical use case. Available: <https://engineering.fb.com/video-engineering/av1-beats-x264-and-libvpxvp9-in-practical-use-case/>.
- [42] L. Guo, J. De Cock, and A. Aaron, "Compression performance comparison of x264, x265, libvpx and aomenc for on-demand adaptive streaming applications," in *2018 IEEE Picture Coding Symposium (PCS)*. IEEE, 2018, pp. 26–30.
- [43] Anthony Vetro,"The MPEG-DASH Standard for Multimedia Streaming Over the Internet", Industry and Standards, Mitsubishi Electric Research Labs, 2011

- [44] Schulzrinne, H., Casner, S., Frederick, R., V. Jacobson, "RTP "A Transport Protocol for Real-Time Applications"
- [45] Αθανάσης Καρούλης, «Διαδυκτιακά πρωτόκολλα ροής (Streaming Protocols) RTP και RTSP», Ελλάδα Ιανουάριος 2001.
- [46] Muller Christopher, "MPEG-DASH (Dynamic Adaptive Streaming over HTTP, ISO/IEC 23009-1)", Bitmovin, 21st April 2015, Available: <https://bitmovin.com/dynamic-adaptive-streaming-http-mpeg-dash/>
- [47] Dash Industry Forum page, Available: <https://dashif.org/>
- [48] M. Covell, M. Arjovsky, Y. chung Lin, and A. Kokaram, "Optimizing transcoder quality targets using a neural network with an embedded bitrate model." Society for Imaging Science and Technology, 2016.
- [49] J.Clement. (2019) Hours of video uploaded to Youtube every minute as of May, 2019. [Online] Available: <https://www.statista.com/statistics/259477/hoursof-video-uploaded-to-youtube-every-minute/>
- [50] G. Esakki, V. Jatla, and M. S. Pattichis, "Optimal hevc encoding based on gop configurations," in 2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI). IEEE, 2016.
- [51] Laderer Stefan, "Why YouTube & Netflix use MPEG-DASH in HTML5", Bitmovin, 2nd February 2015, Available: <https://bitmovin.com/status-mpeg-dash-today-youtube-netflix-use-html5-beyond/>
- [52] Anthony Vetro, "MPEG's Dynamic Adaptive Streaming over HTTP (DASH) – Enabling Formats for Video Streaming over the Open Internet", Industry and Standards, Mitsubishi Electric Research Labs, 2011
- [53] Muller Christopher, "MPEG-DASH (Dynamic Adaptive Streaming over HTTP, ISO/IEC 23009-1)", Bitmovin, 21st April 2015, Available: <https://bitmovin.com/dynamic-adaptive-streaming-http-mpeg-dash/>
- [54] Iraj Sodagar, "Media presentation description and segment formats", Mpeg, Standard: MPEG-DASH, March 2011. Available: <https://mpeg.chiariglione.org/standards/mpeg-dash/media-presentation-description-and-segment-formats>

- [55] MPEG-DASH, Ultrasound Modalities, University of Cyprus, Available:
<http://emcvideoz.cs.ucy.ac.cy/index.html>
- [56] Thomas Stockhammer, “MPEG's Dynamic Adaptive Streaming over HTTP (DASH) – Enabling Formats for Video Streaming over the Open Internet”, Qualcomm Incorporated, Webinarat EBU - Nov 22, 2011
- [57] “TruNET wireless User Manual”, Fractal Networx Ltd, 2018, Available:
<https://www.fractalnetworx.com/>
- [58] “System Architecture, Requirements, and Specifications Report”, ACTRESS, v1.2, 14th June, 2019.