Scriptie 4934 door: Hannes Mareen
Computer Science Engineering

Digitale piraten opsporen met een watermerk

In 2015 werd de totale waarde van illegaal verspreide films en tv-series geschat op 130 miljard euro. De televisiereeks Game of Thrones is daarin een koploper; het breekt al enkele jaren de records van meest illegaal gedownloade tv-programma. Zo werd de laatste aflevering van het zevende seizoen meer dan 120 miljoen keer gedownload in de eerste drie dagen na uitgave! In sommige gevallen gaat het zelf een stap verder en zijn de afleveringen online beschikbaar nog vóór deze officieel uitgebracht worden. Zo werden bijvoorbeeld in 2015 de eerste vier afleveringen van het nieuwe seizoen een dag voor de première gelekt. In één week werden deze maar liefst 32 miljoen keer illegaal gedownload.

De filmindustrie heeft er alle belang bij om digitale piraterij zo snel en efficiënt mogelijk aan te pakken. Sommige organisaties willen dit doen door de (miljoenen) mensen die illegaal downloaden op te sporen en te straffen. Een betere oplossing is echter om het probleem bij de bron aan te pakken, namelijk bij de enkele digitale piraten die de video's illegaal op het internet zetten.

Een mogelijke oplossing is het zogenaamde watermerken van video's. Wanneer een video legaal verstuurd wordt naar een klant, plaatst men in deze video een uniek watermerk dat de identificatiegegevens van deze klant voorstelt. Als een onbetrouwbare klant zijn gewatermerkte versie van de video illegaal op het internet zet, kan men het watermerk uitlezen en de criminele klant identificeren.

Hoe het niet moet: tekst als watermerk

De eerder vermelde afleveringen van Game of Thrones werden op voorhand verstuurd naar vertrouwde recensenten, zodat deze een review klaar zouden hebben op de dag van première. Om de video's te beveiligen kreeg elke recensent een versie met daarin een persoonlijk watermerk; de identificatiegegevens werden weergegeven als tekst in een hoek van de video (zie onderstaande, linkse afbeelding). Dit betekende enerzijds dat alle klanten – ook de eerlijke – gestoord werden. Het watermerk blokkeerde namelijk een deel van de video. Anderzijds was het erg gemakkelijk voor piraten om het watermerk te omzeilen. Ze konden het watermerk simpelweg wazig maken zodat ze niet meer identificeerbaar waren (zie onderstaande, rechtse afbeelding).

Een voorbeeld van hoe men video's niet moet watermerken: het watermerk (links) is storend zichtbaar en gemakkelijk te verwijderen door het wazig te maken (rechts).

Uiteraard bestaan er ook meer geavanceerdere technieken. Deze creëren echter nog altijd geen watermerk dat niet storend zichtbaar is, niet gemakkelijk te verwijderen is én op grote schaal te gebruiken is.

Hoe ik het doe: impliciete aanpassingen als watermerk

In mijn masterproef heb ik een nieuwe techniek ontwikkeld die wél onzichtbaar, niet-verwijderbaar en op

grote schaal te gebruiken is. Wat deze aanpak zo uniek maakt is dat het watermerk automatisch geïntroduceerd wordt door de video-encoder.

Een video-encoder wordt gebruikt om video's te comprimeren. Dit doet men door alle kleine regio's in een video te voorspellen op basis van andere, omringende kleine regio's. Zo is het heel aannemelijk om te voorspellen dat, bijvoorbeeld in een video met planten, naast een bladje van een plant een ander bladje van die plant te zien is. Door dit voorspellend gedrag kan een encoder de bestandsgrootte van een video sterk verkleinen. Bij deze voorspellingen maakt de encoder kleine foutjes, ook wel compressieartefacten genoemd, die men als mens niet opmerkt. Met andere woorden, de encoder comprimeert de video zonder aan kwaliteit in te boeten.

Mijn ontwikkelde watermerkmethode speelt in op dit voorspellend gedrag en de daarbij horende kleine foutjes. Tijdens het comprimeren introduceer ik het watermerk in de video door één kleine regio subtiel aan te passen. Deze aanpassing is zo subtiel dat men dit als mens niet opmerkt, bijvoorbeeld door een klein stukje van een bladje van een plant een andere tint groen te geven. Doordat de video-encoder deze kleine regio gebruikt bij de voorspelling van omliggende regio's, zullen deze omliggende regio's ook (impliciet) aangepast worden. Dit komt doordat de encoder deze omliggende regio's anders zal voorspellen dan wanneer men de kleine regio niet had aangepast. Als gevolg zal de encoder dan ook andere compressieartefacten genereren in deze omliggende regio's. Kortom, door één expliciete aanpassing in één regio worden vele andere regio's ook, impliciet, aangepast.

Doordat elke expliciete aanpassing zich verspreidt tot andere impliciete aanpassingen, stel ik het watermerk voor door deze unieke collectie aanpassingen. Zo kan een bepaald watermerk bijvoorbeeld voorgesteld worden door de impliciete aanpassingen gegenereerd door een bepaalde expliciete aanpassing in de linkerbovenhoek, en een ander watermerk door die gegenereerd door een bepaalde expliciete aanpassing in de rechteronderhoek.

Het grootste voordeel van deze nieuwe techniek is dat de impliciete aanpassingen automatisch gecreëerd worden door de video-encoder en dus normale, veelvoorkomende compressieartefacten zijn. Doordat zo'n kleine foutjes ook voorkomen in video's zonder watermerk zijn mensen reeds gewend om deze constant te zien. Ze worden met andere woorden niet als storend ervaren en worden zelfs niet eens opgemerkt. Een tweede voordeel aan deze techniek is dat een piraat de impliciete aanpassingen niet kan onderscheiden van compressieartefacten die sowieso in de video aanwezig zijn. Hierdoor is het dus ook erg moeilijk – of zelfs onmogelijk – om het watermerk te verwijderen. Ten slotte wordt gebruik gemaakt van slimme encoders die deze kleine expliciete aanpassingen snel kunnen maken, waardoor het watermerksysteem toepasbaar is op relatief grote schaal. Kortom, het watermerk is onzichtbaar, niet-verwijderbaar én op grote schaal te gebruiken.

Mijn excuses

Indien men de eerder vermelde, vroegtijdig gelekte afleveringen van Game of Thrones had beveiligd met mijn watermerktechniek dan had men de piraten kunnen identificeren. Het probleem zou met andere woorden bij de bron aangepakt worden en daardoor vermeden worden in de toekomst. Mijn excuses alvast aan u, de lezer van dit artikel, want u downloadt toch ook wel eens iets illegaal?