

**Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie**

Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki

KATEDRA INFORMATYKI STOSOWANEJ



PRACA MAGISTERSKA

KONRAD MALAWSKI

**PRZETWARZANIE I ANALIZA DANYCH
MULTIMEDIALNYCH W ŚRODOWISKU
ROZPROSZONYM**

PROMOTOR:

dr inż. Sebastian Ernst

Kraków 2013

OŚWIADCZENIE AUTORA PRACY

OŚWIADCZAM, ŚWIADOMY ODPOWIEDZIALNOŚCI KARNEJ ZA
POŚWIADCZENIE NIEPRAWDY, ŻE NINIEJSZĄ PRACĘ DYPLOMOWĄ
WYKONAŁEM OSOBIŚCIE I SAMODZIELNIE, I NIE KORZYSTAŁEM ZE
ŹRÓDEŁ INNYCH NIŻ WYMIENIONE W PRACY.

.....

PODPIS

AGH
University of Science and Technology in Krakow

Faculty of Electrical Engineering, Automatics, Computer Science and
Electronics

DEPARTMENT OF APPLIED COMPUTER SCIENCE



MASTER OF SCIENCE THESIS

KONRAD MALAWSKI

**MEMORY MANAGEMENT IN MOBILE OPERATING
SYSTEMS**

SUPERVISOR:

Sebastian Ernst Ph.D

Krakow 2013

Acknowledgements...

Contents

1. Problem definition..... 7
 1.0.1. Use cases 7
2. Problem definition..... 9
3. Problem definition..... 10
4. Problem definition..... 11
5. Problem definition..... 12

Todo list

1. Problem definition

The goal of this work is to research and development of a solution which would enable matching videos with their copies in lower, or otherwise disturbed quality.

1.0.1. Use cases

One of the simplest use cases in which this system might be used is trivial *duplicate detection* of video files. One might think that a simple checksum of the video files would suffice, but imagine a system like *YouTube* where hours of video content are uploaded each second. It's easy to imagine two people uploading the same trailer for an upcoming movie during the same day. In the easiest case, a simple checksum check would suffice, but what if the videos were encoded using different codecs or even were uploaded in different resolutions?

To further represent how difficult it may be to answer the question of „Are those two files the same material?“ let's go back to the example of users uploading video content to our YouTube-like service. What if the users knew the uploaded material is copy righted and should not be uploaded to the public internet? As one might imagine, this might happen quite often on a service like YouTube. Let's assume our users know about our „duplicate“ detection algorithm and that we might be partnering with movie companies, so that we have reference material on the servers we can compare the uploaded videos to. Obviously the users will try to trick our algorithm into not recognising that the uploaded content is the same as our reference material. One of the tricks often seen on YouTube.com is that the users upload the material „mirrored“ which further complicates our matching algorithm – we must now also be resilient against data modified especially for the goal of us not being able to detect it.

Another, less copyright focused, goal of the presented system is to be able to extensively mine data from the uploaded video content. Here the canonical example would be a „*TOP 10 Movies of All Time*“ video, which obviously contains video material from at least 10 movies, usually in the order of 10th, 9th ... until the 1st (best) movie of all time. If we would be able to match parts of each video to their corresponding reference materials, we would be able to get meta data about the now recognised movies and even mine out the data what is the best / worst movie of all time, even without it being written per se - only by looking at the frames in the

video. This idea only scratches the surface of what the system implemented during this thesis work might do, but it will be our test case that we'll be working towards during this paper.

The system implemented during this work represents a basic effort to tackle the above problems with the goal of being able to identify

2. Problem definition

3. Problem definition

4. Problem definition

5. Problem definition