

# Graph database watermarking using pseudo nodes

## 1. Introduction

- Media is constantly transferred over an open channel, which allows for high amount of unauthorized or fake data to be marketed as real and reliable in the open
- Watermarking provides a way to prove authenticity and integrity of the data

## 3. Terminology

- ``_id` field` - The field name `_id` is reserved for use as a primary key
- **Pseudo document** – a fake document designed to hold the watermark data

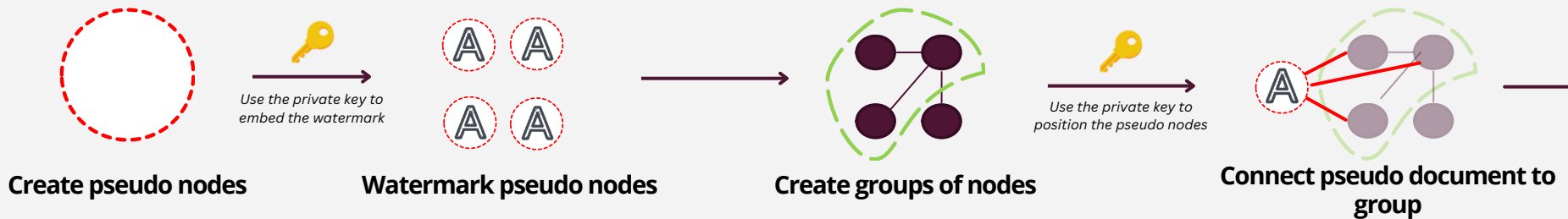
## 2. Research

The research introduces an improved method for watermarking by identifying and improving upon flaws inside already established watermarking technique. It focuses on a new watermarking technique introduced in 2022, studying it in detail and identifying possible ways the algorithm can be broken.

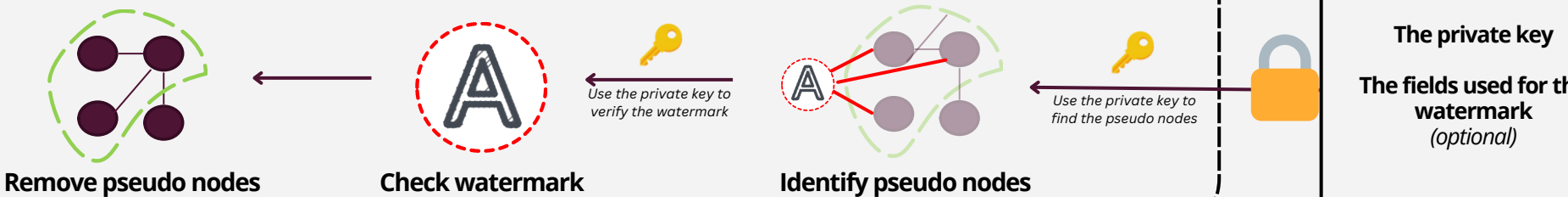
## 4. Research

A literature study is performed to find relevant watermarking techniques on NoSQL databases. A technique is chosen and analyzed. Then, improvements to the algorithm are implemented and tested to determine how the applied changes influenced the security of the watermark.

## Embed watermark



## Verifying watermark

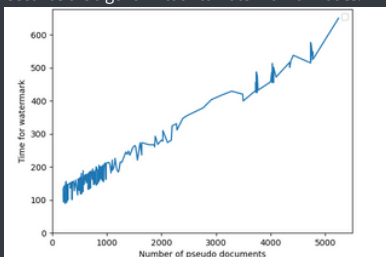


## Secret

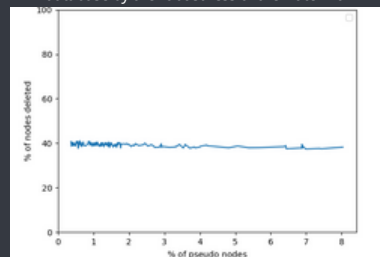
The private key  
The fields used for the watermark (optional)

## 5. Analysis

The algorithm's performance is measured in the number of seconds the algorithm took to watermark all nodes:



Usability is measured by ratio of pseudo nodes inside the database by the robustness of the watermark



## 6. Conclusion

We introduce a new graphical database watermarking algorithm based on a novel approach of inserting watermarked pseudo nodes. This approach promises to create a robust, non-detectable, and secure watermark, resilient to modification, deletion and insertion attacks. A new algorithm is designed and implemented to tackle flaws in previous algorithms and then tested against modification and deletion attacks. The test results showed that the algorithm underperformed in contrast to the algorithm it inspires from. This means that further research needs to be done for a robust and secure algorithm to be introduced for graph database watermarking.

Author: **Tsvetomir Hristov** (T.Hristov@student.tudelft.nl)  
Supervisors: **Dr. Zekeriya Erkin** and **Devris Isler**

References:  
Erkin, Dr Zeki. 'Dataset Watermarking'. Project Forum, 2022. [https://projectforum.tudelft.nl/course\\_editions/5/2/generic\\_projects/1b2](https://projectforum.tudelft.nl/course_editions/5/2/generic_projects/1b2).

Zhuang, Xiaodan, et al. 'Document Database Watermark Algorithm Based on Connected Graph'. 2022 International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT), 2022, pp. 212-218. <https://doi.org/10.1109/CCPQT56151.2022.00044>.

Kumar, Sanjay, et al. 'A Recent Survey on Multimedia and Database Watermarking'. Multimed. Tools Appl., vol. 79, no. 27-28, 2020, pp. 20149-20197. <https://doi.org/10.1007/s11042-020-08881-y>.