Database Watermarking: Protection by Alteration

What are the benefits from database watermarking techniques for DB preservation purposes

David Gross-Amblard

IM team – Le2i Lab.
Bourgogne University, Dijon, France
http://ufrsciencesetech.u-bourgogne.fr/~gadavid

Vertigo team - Cedric Lab.
CNAM Paris, France

Funded by ACI Sécurité & Informatique Tadorne 2004-2007
Outline

1. What are the Threats on Databases Preservation Systems?

2. Watermarking: Pervasive Information Insertion

3. Questions raised by a Watermarking Approach
DB Preservation Goals

“Any serious long-term preservation strategy for any kind of digital content pursues to guarantee continuously:

- **Integrity**: protected from unintended and intended harm
- **Intelligibility**: understandable and comprehensible
- **Authenticity**: authentication (of authorship and provenance) and reliability (of the contained evidence / accuracy)
- **Originality**: as close to the original as possible
- **Accessibility**: technically readable and usable to users”

[peter Keller-Marxer, Swiss federal archives]

- **Accessibility**: technically readable and usable to authorized users (redaction of restricted content)
DB Preservation Goals

“Any serious long-term preservation strategy for any kind of digital content pursues to guarantee continuously:

- **Integrity**: protected from unintended and intended harm
- **Intelligibility**: understandable and comprehensible
- **Authenticity**:
  - authentication (of authorship and provenance)
  - and reliability (of the contained evidence / accuracy)
- **Originality**: as close to the original as possible
- **Accessibility**: technically readable and usable to users’

[peter Keller-Marxer, Swiss federal archives]

- **Accessibility**: technically readable and usable to **authorized** users
  (redaction of restricted content)
Threats on (Physical or Digital) Archives

Threats

- **Neglect**
- **Vandalism** (usually revenge)
- **Theft** (or information leakage)
- **Storage destruction**

[2004 Nationaal Archief (the National Archives of the Netherlands) and the European Commission on Preservation and Access]

**New context**

- One official data producer, one preservation agency, and nice users
- Many production and preservation locations
  - different places storage recommendation
- Few physical objects, difficult to “create”
- Huge stream of purely digital objects, easy to forge and copy
Threats, Role by Role, and new needs

- **Bad data producer:**
  - try to submit *fake data* (e.g. not produced by the correct device)
  - check that recorded data are authentic, trace information leakage (acting as a simple user)

- **Bad archivist:**
  - alter data by e.g. format mismatch (neglect), or voluntarily
  - add *fake records*
  - export records with restricted access
  - check authenticity prior to recording, trace information leakage

- **Bad user:**
  - use records to make a *fake archive* (and information leakage)
  - check authenticity of the archive site, find the authentic archive from redacted excerpts
Threats, Role by Role , and new needs

- **Bad data producer :**
  - try to submit **fake data** (e.g. not produced by the correct device)
  - check that recorded data are authentic, trace information leakage
    (acting as a simple user)

- **Bad archivist :**
  - alter data by e.g. format mismatch (neglect), or voluntarily
  - add **fake records**
  - export records with **restricted access**
  - check authenticity prior to recording, trace information leakage

- **Bad user :**
  - use records to make a **fake archive** (and information leakage)
  - check authenticity of the archive site, find the authentic archive from
    redacted excerpts
Analogy: Ancient Greek and Medieval Texts Archives
Analogy: The Greek Watermark Archive

"Its purpose is to facilitate:

- **dating** of Greek manuscripts
- establishing the **provenance** of Greek manuscripts
- identifying **scribes** of unsigned Greek manuscripts
- reconstructing the productivity of later Byzantine and post-Byzantine scribes and centers of Greek book production from the evidence of paper
- research in the history of paper production and use"

[The Watermark Archive / Maine & Thessaloniki]
Digital Watermarking: Pervasive Information Insertion

- Information hiding by controlled data alteration
- Robust watermarking: resilient to malevolent data transformations
- Trade-off reliability of data/robustness of the watermark

![Diagram of watermark insertion and extraction process]

- **Original record**
- **Watermark**
- **Watermarked record**

- **Extractor**
  - **Message**
  - **Suspect record**

- **Marker**
  - **Secret key**
  - **Message**

Gross-Amblard (IM-LE2I/Vertigo-CNAM)
Q: “How can we Preserve Authenticity and Provenance of Databases?”

Bad Archivist:

- Onerous nuclear CERN experiment, huge amount of data
- Several archivists (several copies) $DB_1, \ldots, DB_k$
- One is selling data prior to official publication

How to trace the bad archivist?

- Add a distinct mark $M_1, \ldots, M_k$, to each recorded db
- On suspect data $DB$, apply the extractor $\rightarrow M_i$
- Bad archivist: number $i$
Q: “How can we Preserve Authenticity and Provenance of Databases?”

Bad Archivist:

- Onerous nuclear CERN experiment, huge amount of data
- Several archivists (several copies) $DB_1, \ldots, DB_k$
- One is selling data prior to official publication

How to trace the bad archivist?

- Add a distinct mark $M_1, \ldots, M_k$, to each recorded db
- On suspect data $DB$, apply the extractor $\rightarrow M_i$
- Bad archivist: number $i$
Example of Relational Database Watermarking
[Agrawal, Haas and Kiernan scheme, VLDB’02]

primary key alterable data

secret key $K$
hash function

which tuple to mark?
$hash(P,K) \mod rate = 0$?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.13957018</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.1349766</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.54730</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.7711</td>
<td>rho(770)</td>
<td>0, +</td>
</tr>
<tr>
<td>223</td>
<td>0.78257</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.95778</td>
<td>eta’(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.980</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.9847</td>
<td>a(0)(980)</td>
<td>0, +</td>
</tr>
</tbody>
</table>

[Particle Data Group]
Example of Relational Database Watermarking

primary key alterable data

secret key $\mathcal{K}$

hash function

which tuple to mark?

$\text{hash}(P.\mathcal{K}) \mod \text{rate} = 0$?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.13957018</td>
<td>$\pi$</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.1349766</td>
<td>$\pi$</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.54730</td>
<td>$\eta$</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.7711</td>
<td>$\rho(770)$</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.78257</td>
<td>$\omega(782)$</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.95778</td>
<td>$\eta'(958)$</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.980</td>
<td>$f(0)(980)$</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.9847</td>
<td>$a(0)(980)$</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Example of Relational Database Watermarking

primary key alterable data

secret key $\mathcal{K}$
hash function
which tuple to mark?
$\text{hash}(P.\mathcal{K}) \mod \text{rate} = 0$?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.13957018</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.1349766</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.54730</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.7711</td>
<td>rho(770)</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.78257</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.95778</td>
<td>eta'(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.980</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.9847</td>
<td>a(0)(980)</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Example of Relational Database Watermarking

primary key alterable data
\( P \) (viewed in binary)

secret key \( \mathcal{K} \)
hash function

which bit to mark?
\[ \text{hash}(P, \mathcal{K}) \mod \text{distortion} = ? \]

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.0101101</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.110101011</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.01011010</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.11101011</td>
<td>rho(770)</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.1001010</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.1111101</td>
<td>eta’(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.111101</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.11111001</td>
<td>a(0)(980)</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Example of Relational Database Watermarking

primary key alterable data

secret key $K$ newline hash function

hide the mark

$\text{hash}(P.K)$

mod 2

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.010110X0</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.110101011</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.01011010</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.11101011</td>
<td>rho(770)</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.10010X10</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.1111101</td>
<td>eta'(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.111101</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.1111100X0</td>
<td>a(0)(980)</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Watermark Extraction on Suspect Data

secret key $\mathcal{K}$

locate tuples

$\text{hash}(P.\mathcal{K}) \mod \text{rate} = 0$?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.13957018</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.1349766</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.54730</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.7711</td>
<td>rho(770)</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.78257</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.95778</td>
<td>eta'(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.980</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.9847</td>
<td>a(0)(980)</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Watermark Extraction on Suspect Data

secret key $\mathcal{K}$

which bit was marked

$\text{hash}(P.\mathcal{K})$ mod distortion = ?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.010110 0</td>
<td>pi</td>
<td>+</td>
</tr>
<tr>
<td>111</td>
<td>0.110101011</td>
<td>pi</td>
<td>0</td>
</tr>
<tr>
<td>221</td>
<td>0.01011010</td>
<td>eta</td>
<td>0</td>
</tr>
<tr>
<td>113</td>
<td>0.11101011</td>
<td>rho(770)</td>
<td>0,+</td>
</tr>
<tr>
<td>223</td>
<td>0.10010 1 0</td>
<td>omega(782)</td>
<td>0</td>
</tr>
<tr>
<td>331</td>
<td>0.1111101</td>
<td>eta’(958)</td>
<td>0</td>
</tr>
<tr>
<td>9010221</td>
<td>0.111101</td>
<td>f(0)(980)</td>
<td>0</td>
</tr>
<tr>
<td>900001</td>
<td>0.1111100 1</td>
<td>a(0)(980)</td>
<td>0,+</td>
</tr>
</tbody>
</table>
Watermark Extraction on Suspect Data

secret key $\mathcal{K}$

does hidden bit corresponds to $\text{hash}(P.\mathcal{K})$ mod 2 ?

<table>
<thead>
<tr>
<th>particle id</th>
<th>value (GeV)</th>
<th>name</th>
<th>charge</th>
<th>charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td>0.0101100</td>
<td>pi</td>
<td>+</td>
<td>NO</td>
</tr>
<tr>
<td>111</td>
<td>0.110101011</td>
<td>pi</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>221</td>
<td>0.01011010</td>
<td>eta</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>0.11101011</td>
<td>rho(770)</td>
<td>0,+</td>
<td></td>
</tr>
<tr>
<td>223</td>
<td>0.1001010</td>
<td>omega(782)</td>
<td>0</td>
<td>YES</td>
</tr>
<tr>
<td>331</td>
<td>0.1111101</td>
<td>eta'(958)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9010221</td>
<td>0.111101</td>
<td>f(0)(980)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>900001</td>
<td>0.11111001</td>
<td>a(0)(980)</td>
<td>0,+</td>
<td>YES</td>
</tr>
</tbody>
</table>

2/3 of matching ratio : suspect !
State of the Art for DB watermarking

- Apply on numerical data / Accuracy-preserving
- Blindness: no need of the original at extraction
- Robustness: resist to voluntarily alteration of data
- Capacity:
  - 1 bit: [Agrawal, Haas, Kiernan, VLDB’02]
  - message: [Sion et al., SIGMOD’03], [Li et al., TDSC’05]

Query-preserving Watermarking
- greedy search [Sion, et al., SIGMOD’03]
- capacity lower bound for FO and MSO [Gross-Amblard, PODS’03],
- optimized search [Lafaye et al., subm.]

Software:
- Watermill, query-preserving watermarking engine [Lafaye et al.]
- controlled alteration using constraints:
  - distortion 0.1 on CERNDatabase.energy
  - distortion 0 on sum(CERNDatabase.energy) where pid<100
And also

- Collusion secure: [Li et al., TDSC’05], [Lafaye et al., subm.]
- Streams: numerical [Sion, et al., TKDE’06], XML [Lafaye et al., DBSec’06]
- Spatial databases [Lafaye et al., subm.], [Ohbuchi et al., SMI’03], etc.
- Fragile watermarks [Guo et al., Inf. Sci.’06]
- Reversible watermarks [Zhang et al., JOC’2006]
- Categorical data [Sion et al. TKDE’05]
Questions raised by a Watermarking Approach

Q :”Can we move from a centralised model to a distributed, redundant model of database preservation?”
  ▶ Who are we going to trust, to what extent? (producer, archivist, user)

Q : “What are the salient features of a database that should be preserved?”

→ If alterations are to be performed for watermarking, to what extent?
  ▶ Don’t touch a single bit!
  ▶ Accuracy of numerical data
  ▶ Accuracy of texts?
  ▶ Do what you what, but don’t touch the views (user views)
    ★ Views over time?
    ★ New analysis of old data

Thanks.
Questions raised by a Watermarking Approach

- Q:”Can we move from a centralised model to a distributed, redundant model of database preservation?”
  - Who are we going to trust, to what extent? (producer, archivist, user)

- Q: “What are the salient features of a database that should be preserved?”

- If alterations are to be performed for watermarking, to what extent?
  - Don’t touch a single bit!
  - Accuracy of numerical data
  - Accuracy of texts?
  - Do what you what, but don’t touch the views (user views)
    - Views over time?
    - New analysis of old data

Thanks.
Questions raised by a Watermarking Approach

Q :”Can we move from a centralised model to a distributed, redundant model of database preservation?”
  ▶ Who are we going to trust, to what extent? (producer, archivist, user)

Q : “What are the salient features of a database that should be preserved?”
  → If alterations are to be performed for watermarking, to what extent?
    ▶ Don’t touch a single bit!
    ▶ Accuracy of numerical data
    ▶ Accuracy of texts?
    ▶ Do what you what, but don’t touch the views (user views)
      ★ Views over time?
      ★ New analysis of old data

Thanks.
Questions raised by a Watermarking Approach

- Q : ”Can we move from a centralised model to a distributed, redundant model of database preservation?”
  - Who are we going to trust, to what extent? (producer, archivist, user)
- Q : “What are the salient features of a database that should be preserved?”
- If alterations are to be performed for watermarking, to what extent?
  - Don’t touch a single bit!
  - Accuracy of numerical data
  - Accuracy of texts?
  - Do what you want, but don’t touch the views (user views)
    - Views over time?
    - New analysis of old data

Thanks.
Questions raised by a Watermarking Approach

- Q :”Can we move from a centralised model to a distributed, redundant model of database preservation?”
  - Who are we going to trust, to what extent? (producer, archivist, user)
- Q : “What are the salient features of a database that should be preserved?”
- → If alterations are to be performed for watermarking, to what extent?
  - Don’t touch a single bit!
  - Accuracy of numerical data
  - Accuracy of texts?
  - Do what you what, but don’t touch the views (user views)
    - Views over time?
    - New analysis of old data

Thanks.