Improving the Imperceptibility of Curve Fingerprinting

MERIT 2004

Dan Cardy / Hongmei Gou and Min Wu

Fingerprinting

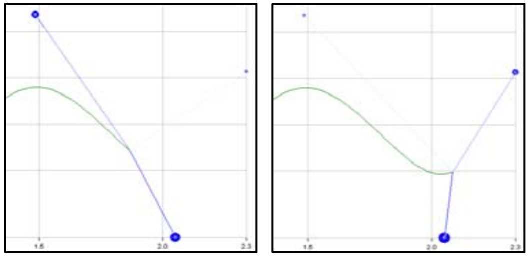
- **Why:** Many documents contain sensitive information that must be distributed to allies, but hidden from foes. If a leak occurs, fingerprinting allows the source to be identified.
- **How:** A fingerprinted document contains hidden information that identifies its intended recipient. This information needs to be:
  - Robust – resistant to tampering.
  - Imperceptible – invisible to the eye.

Curve Fingerprinting

- Many documents are primarily composed of curves:
  - Handwritten text and hand-drawn images
  - Maps
  - CAD Drawings
- We developed a method for embedding information into curves by slightly changing the shape of the curves.
- This technique resists noise and these attacks:
  - Collusion
  - Geometric Transformations
  - Printing and Scanning

B-splines

B-splines are piecewise polynomial functions used to approximate curves from a small set of characterizing data called control points.

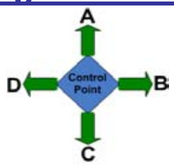


A weighted average of three control points determines each point along the curve. The weights change as you move along the curve.

Curve Fingerprinting with B-splines

➢ To create a marked document for distribution:

1) Create a B-spline representation of the original curve; reduce the original to a set of control points.
2) Embed a fingerprint by slightly moving the control points according to a noise-like sequence unique to each user.

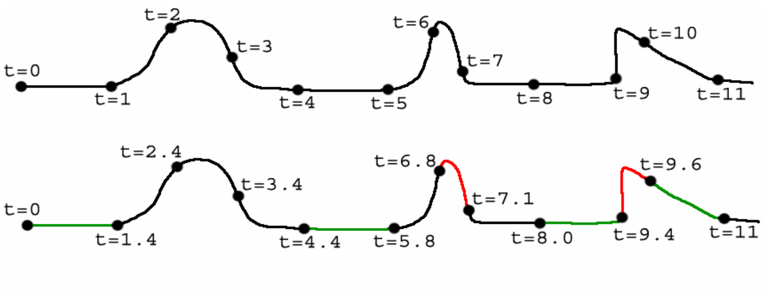


➢ To trace the origin of a leaked copy:

1) Reduce it to a set of control points.
2) Compare with the original control points to extract the fingerprint.

Imperceptibility of Curve Fingerprinting

- Depends the strength of fingerprint.
- Depends on the goodness of B-spline approximation. If the B-spline representation differs noticeably from the original curve, this distortion will be exacerbated by the distortion added by the embedding process.
- A new algorithm is developed to address the B-spline approximation issue.



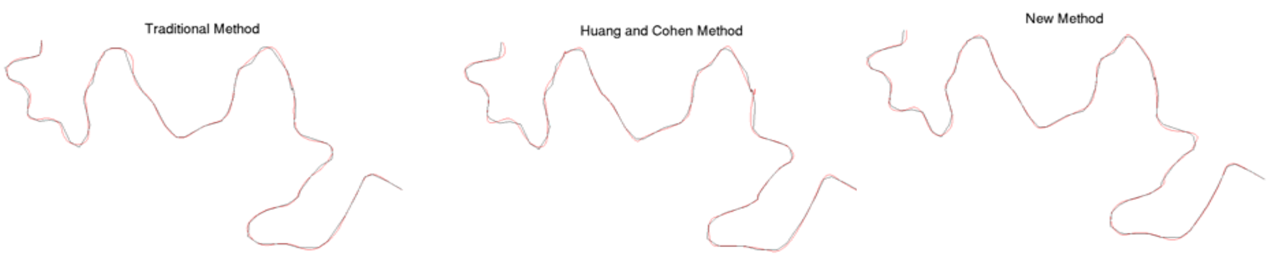
The New Algorithm

- When a curve is reduced to n control points, this system of equations must be solved:

  B*(n control points) = (curve points)

- B contains only the k values of the B-spline $B_k(t)$ for some $t \in [0,n-1]$. One t value is associated with each point on the curve.
- By manipulating the t values assigned to a group of points, we manipulate the importance of placing a control point near that group.
  - Traditionally, t values increase proportionally with arc length.
  - Huang and Cohen improve on this by increasing t in inverse proportion to the arc length.
  - The new algorithm increases t in inverse proportion to the curvature at each point, which means more control points are allocated to high-curvature segments.

Results



The original curve is in black. The approximated curves are in red. The new algorithm generates the least noticeable approximation errors.