Asymptotic fingerprinting capacity in the combined digit model

Asymptotic fingerprinting capacity in the Combined Digit Model

Dion Boesten Boris Škorić

Eindhoven University of Technology

{d.boesten, b.skoric}@tue.nl

Abstract

We study the channel capacity of q-ary fingerprinting in the limit of large attacker coalitions. We extend known results by considering the Combined Digit Model, an attacker model that captures signal processing attacks such as averaging and noise addition. For $q = 2$ we give results for various attack parameter settings.

1 Introduction

Watermarking is a means of tracing the (re-)distribution of content. Before distribution, digital content is modified by applying an imperceptible watermark (WM). Once an unauthorized copy of the content is found, the WM helps to trace those users who participated in the creation of the copy. Reliable tracing requires resilience against attacks that aim to remove the WM. Collusion attacks are a particular threat: multiple users cooperate, and differences between their versions of the content tell them where the WM is located. Coding theory has provided a number of collusion-resistant codes. The resulting system has two layers: The coding layer determines which message to embed, and protects against collusion attacks. The underlying watermarking layer hides symbols of the code in segments of the content. Many codes have been proposed in the literature. Most notable is the Tardos code [16], which achieves the asymptotically optimal proportionality $m \propto c^2$, with m the code length and c the size of the coalition. Tardos introduced a two-step stochastic procedure for generating codewords: (i) For each segment a bias is randomly drawn. (ii) For each user independently, a 0 or 1 is drawn for each segment using the bias for that segment. This construction was generalized to larger (q-ary) alphabets in [17].
The interface between the coding and WM layer is specified in terms of the *Marking Assumption* (MA), which states that the colluders can attack only in those segments where they received different WM symbols. These are called detectable positions. There is a further classification of attacks according to the manipulations that can be performed *in the detectable positions*. In the *Restricted Digit Model* (RDM), the coalition is only allowed to pick one symbol that they received. In the *Unreadable Digit Model* (UDM), they are further allowed to create an erasure. In the *Arbitrary Digit Model*, they can pick any symbol, even one that they did not receive. The *General Digit Model* allows any symbol or an erasure. For $q = 2$, all these MA attacks are equivalent. For $q > 2$, the general feeling is that realistic attacks are somewhere between the RDM and the UDM. To get an even more realistic attack model which takes into account signal processing (e.g. averaging attacks and noise addition), one has to depart from the MA. Such models were proposed in [20] and [18] for general q, and for $q = 2$ in e.g. [8, 9].
In Tardos' scheme [16] and later improvements (e.g. [19, 17, 3, 15, 14, 5, 18, 20, 11, 10, 12]), users are found to be innocent or guilty via an 'accusation sum', a sum of weighted per-segment contributions, computed for each user separately. The analysis of achievable performance was helped by an information-theoretic treatment of anti-collusion

codes. Bias-based codes can be treated as a maximin game [2, 13, 7], independently played for each segment, where the payoff function is the mutual information between the symbols $x_1, \ldots, x_c$ handed to the colluders and the symbol y produced by them. In each segment the colluders try to minimize the payoff function using an attack strategy that depends on the received symbols $x_1, \ldots, x_c$. The watermarker tries to maximize the payoff by setting the bias distribution.

The rate of a fingerprinting code is defined as $(\log_q n)/m$, with n the number of users and m the code length. The *fingerprinting capacity* is the maximum achievable rate. For $q = 2$ it was conjectured [7] that the capacity is asymptotically $1/(c^2 2 \ln 2)$. The conjecture was proved in [1, 6]. In [1] an accusation scheme was developed where candidate coalitions get a score related to the mutual information between their symbols and y. It achieves capacity but is computationally too expensive. Huang and Moulin [6] proved for the large-c limit (for $q = 2$) that the interleaving attack and Tardos's arcsine distribution are optimal. It was shown in [4] that the asymptotic channel capacity for q-ary alphabets in the RDM is $(q-1)/(2c^2 \ln q)$.

In this paper we study the asymptotic fingerprinting capacity in the Combined Digit Model (CDM) [18]. We choose for the CDM because this model is defined for general q and captures a range of non-MA attacks. We show that the asymptotic channel capacity in the CDM can be found by solving the following problem: Find a mapping γ from the hypersphere in q dimensions to the hypersphere in 2^q dimensions, such that γ minimizes the volume swept in the latter space; the boundary conditions on the volume are fixed by the parameters in the CDM. For $q \geq 3$ we have not solved the minimization problem. For $q = 2$ we present numerical results. The numerics involve computations of constrained geodesics, a difficult problem in general. The resulting graphs show a nontrivial dependence of the capacity on the CDM attack parameters.

2 Preliminaries

2.1 Fingerprinting with per-segment symbol biases

We use capital letters for random variables, and lowercase letters for their realizations. Vectors are in boldface and the components of a vector $\vec{x}$ are written as x_i. Vectors are interpreted as being column vectors. The expectation over X is denoted as $\mathbb{E}_X$. The mutual information between X and Y is denoted by $I(X;Y)$, and the mutual information conditioned on a third variable Z by $I(X;Y|Z)$. The base-q logarithm is written as $\log_q$. The standard Euclidean norm of a vector $\vec{x}$ is denoted by $\|\vec{x}\|$.

Tardos [16] introduced the first fingerprinting scheme that achieves optimality in the sense of having the asymptotic behavior $m \propto c^2$. He introduced a two-step stochastic procedure for generating the codeword matrix X. Here we show the generalization to non-binary alphabets [17]. A Tardos code of length m for a number of users n over the alphabet $\mathcal{Q}$ of size q is an $n \times m$ matrix of symbols from $\mathcal{Q}$. The codeword for user i is the i'th row in X. An auxiliary bias vector $\vec{P}^{(j)} \in [0,1]^q$ with $\sum_\alpha P_\alpha^{(j)} = 1$ is generated independently for each column j, from a distribution F which is considered known to the attackers. Each entry X_{ij} is generated independently: $\text{Prob}\,[X_{ij} = \alpha] = p_\alpha^{(j)}$.

2.2 The Combined Digit Model

Let the random variable $\Sigma_\alpha^{(j)} \in \{0, 1, \ldots, c\}$ denote the number of colluders who receive the symbol α in segment j. It holds that $\sum_\alpha \sigma_\alpha^{(j)} = c$ for all j. From now on we will drop the segment index j, since all segments are independent. In the *Restricted Digit Model* the colluders produce a symbol $Y \in \mathcal{Q}$ that they have seen at least once. In

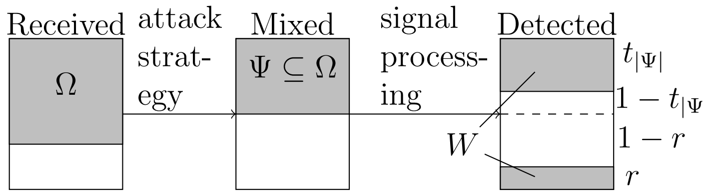


Figure 1: *Overview of the Combined Digit Model.*

the *Combined Digit Model* as introduced by [18] we also allow the attackers to output a mixture of symbols. Let $\Omega(\Sigma) \triangleq \{\alpha \in \mathcal{Q} \mid \Sigma_\alpha \geq 1\}$ be the set of symbols that the pirates have seen in a certain column. Then the output of the pirates is a non-empty set $\Psi \subseteq \Omega(\Sigma)$. On the watermarking level this represents a content-averaging attack where all symbols in Ψ are used. It is sufficient to consider a probabilistic per-segment (column) attack which does not distinguish between the different colluders. Such an attack then only depends on $\vec{\Sigma}$, and the strategy can be completely described by a set of probabilities $\theta_{\psi|\vec{\sigma}} \in [0, 1]$, which are defined as $\theta_{\psi|\vec{\sigma}} \triangleq \text{Prob}[\Psi = \psi \mid \vec{\Sigma} = \vec{\sigma}]$.
The CDM also introduces a stochastic detection process. Let $|\Psi|$ be the cardinality of the output set Ψ. Then each symbol in Ψ is detected with probability $t_{|\Psi|}$. Each symbol not in Ψ is detected with error probability r. The set $W \subseteq \mathcal{Q}$ indicates which symbols are detected. Note that Ψ is forced to be non-empty, but $W = \emptyset$ can occur.
The numbers t_i are decreasing since mixing more symbols makes it more difficult to detect the individual symbols. The overall probability of detecting a set w, given ψ, is

$$M_{w|\psi} = t_{|\psi|}^{|w\cap\psi|} \left(1 - t_{|\psi|}\right)^{|\psi\setminus w|} r^{|w\setminus\psi|} (1 - r)^{q-|w\cup\psi|} . \tag{1}$$

These probabilities form a $2^q \times (2^q - 1)$ matrix M. In this way we can define

$$\tau_{w|\vec{\sigma}} \triangleq \text{Prob}\left[W = w \mid \vec{\Sigma} = \vec{\sigma}\right] = \sum_{\psi} M_{w|\psi}\theta_{\psi|\vec{\sigma}} = (M\theta)_{w|\vec{\sigma}} . \tag{2}$$

2.3 Collusion channel and fingerprinting capacity

Similarly to the RDM [4] the attack can be interpreted as a noisy channel with input $\vec{\Sigma}$ and output W. A capacity for this channel can then be defined, which gives an upper bound on the achievable code rate of a reliable fingerprinting scheme. The first step of the code generation, drawing the biases $\vec{p}$, is not considered to be a part of the channel. The fingerprinting capacity C_q^{CDM} for a coalition of size c and alphabet size q in the CDM is equal to the optimal value of the following two-player game:

$$C_q^{\text{CDM}} = \max_F \min_{\vec{\theta}} \frac{1}{c} I(W; \vec{\Sigma} \mid \vec{P}) = \max_F \min_{\vec{\theta}} \frac{1}{c} \int F(\vec{p}) I(W; \vec{\Sigma} \mid \vec{P} = \vec{p}) \mathrm{d}^q\vec{p}. \tag{3}$$

Here the information is measured in q-ary symbols. Our aim is to compute the fingerprinting capacity C_q^{CDM} in the limit $(n \to \infty, c \to \infty)$. The payoff function $I(W; \vec{\Sigma} \mid \vec{P})$ is linear in F and convex in $\vec{\tau}$. Because $\vec{\tau} = M\vec{\theta}$ is linear in $\vec{\theta}$ the game is also convex in $\vec{\theta}$ and we can apply Sion's Theorem:

$$\max_F \min_{\vec{\theta}} I(W; \vec{\Sigma} \mid \vec{P}) = \min_{\vec{\theta}} \max_F \ I(W; \vec{\Sigma} \mid \vec{P}) = \min_{\vec{\theta}} \max_p \ I(W; \vec{\Sigma} \mid \vec{P} = \vec{p}), \tag{4}$$

where we did the maximization over F by choosing the optimum $F^*(\vec{p}) = \delta(\vec{p} - \vec{p}_{\max})$ at the location $\vec{p} = \vec{p}_{\max}$ of the maximum of $I(W; \vec{\Sigma} \mid \vec{P} = \vec{p})$.

3 Asymptotic analysis for general alphabet size

We are interested in how the payoff function $I(W;\Sigma \mid \vec{P}=\vec{p})$ of the alternative game (4) behaves as c goes to infinity. Following the same approach as in [4] our starting point is the observation that the random variable $\vec{\Sigma}/c$ tends to a continuum in $[0,1]^q$ with mean $\vec{p}$. We introduce the following notation:

$$h_\psi(\vec{\sigma}/c) \quad \overset{c\to\infty}{=} \quad \theta_{\psi|\vec{\sigma}}. \tag{5}$$

$$g_w(\vec{\sigma}/c) \quad \overset{c\to\infty}{=} \quad \tau_{w|\vec{\sigma}} = \textstyle\sum_\psi M_{w|\psi} h_\psi(\vec{\sigma}/c), \tag{6}$$

which can be written as $\vec{g}=M\vec{h}$. Next we do a 2nd order Taylor expansion of $g_w\left(\frac{\vec{\sigma}}{c}\right)$ around the point $\frac{\vec{\sigma}}{c}=\vec{p}$. This allows us to expand I in powers of $1/c$, giving (see [4])

$$I(W;\Sigma \mid \vec{P}=\vec{p}) = \frac{T(\vec{p})}{2c\ln q} + \mathcal{O}(c^{-3/2}) \tag{7}$$

$$T(\vec{p}) \triangleq \textstyle\sum_w \frac{1}{g_w(\vec{p})} \sum_{\alpha\beta} K_{\alpha\beta} \frac{\partial g_w(\vec{p})}{\partial p_\alpha}\frac{\partial g_w(\vec{p})}{\partial p_\beta}, \tag{8}$$

where $K_{\alpha\beta}=\delta_{\alpha\beta}p_\alpha - p_\alpha p_\beta$ is the scaled covariance matrix of Σ. The capacity $C^{\rm CDM}_{q,\infty}$ in the limit of $c\to\infty$ is then the solution of the continuous version of the game (4):

$$C^{\rm CDM}_{q,\infty} \triangleq \frac{1}{2c^2\ln q} \min_{\vec{h}} \max_{\vec{p}} T(\vec{p}). \tag{9}$$

We introduce variables $u_\alpha \triangleq \sqrt{p_\alpha}$, $\gamma_w \triangleq \sqrt{g_w}$ and the $2^q\times q$ Jacobian matrix $J_{w\alpha}(\vec{u}) \triangleq \frac{\partial\gamma_w(\vec{u})}{\partial u_\alpha}$. We switch to hyperspheres ($\|\vec{u}\|=1, \|\gamma\|=1$) instead of the hyperplanes ($\sum_\alpha p_\alpha=1, \sum_w g_w=1$). The function $\vec{\gamma}(\vec{u})$ was originally defined only on $\|\vec{u}\|=1$, but the Taylor-expansion forces us to define it on a larger domain, i.e. slightly away from $\|\vec{u}\|=1$. There are many consistent ways to do this. We define $\vec{\gamma}$ independent of the radial coordinate $\|\vec{u}\|$. This yields $J\vec{u}=0$, which allows us to simplify $T(\vec{u})$ to

$$T(\vec{u}) = \textstyle\sum_{w,\alpha}(\partial\gamma_w/\partial u_\alpha)^2 = \mathrm{Tr}(J^TJ) = \sum_{i=1}^{q-1}\lambda_i(\vec{u}), \tag{10}$$

where $\lambda_i(\vec{u})$ are the eigenvalues of J^TJ. Because of $J\vec{u}=0$ we know that one of the eigenvalues is 0 with eigenvector $\vec{u}$. Hence $i\in\{1,\ldots,q-1\}$. We wish to find $\min_\gamma \max_u T(u)$ under the constraint $\gamma_w=\sqrt{g_w}=\sqrt{(Mh)_w}$, with M known and

$$h_\psi \geq 0 \quad \forall\psi, \qquad\qquad \textstyle\sum_\psi h_\psi = 1. \tag{11}$$

The constraint $g=Mh$ makes the min-max game more difficult. It is not possible to use the same machinery as for the RDM. For $q=2$ we are however able to compute the asymptotic capacity.

4 Fingerprinting capacity in the CDM for $q=2$

4.1 Solving the max-min game

For $q=2$ the expression (10) simplifies to $T(\vec{u})=\mathrm{Tr}(J^TJ)=\lambda(\vec{u})$ since there is only one nonzero eigenvalue. Furthermore we have the relation $\mathrm{d}\vec{\gamma}=J\mathrm{d}\vec{u}$ and $\|\mathrm{d}\vec{\gamma}\|=$

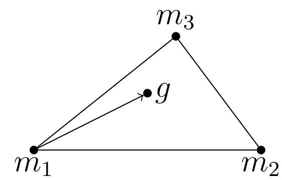


Figure 2: *The vector $\vec{g}$ is not allowed to lie outside the triangle.*

$\sqrt{\lambda}\|\mathrm{d}\vec{u}\|$. We proceed by rewriting

$$\max_{\vec{u}} T(\vec{u}) = \max_{\vec{u}} \lambda(\vec{u}) = \left(\max_{\vec{u}} \sqrt{\lambda(\vec{u})}\right)^2$$
$$\geq \left(\frac{\int \sqrt{\lambda(\vec{u})}\|\mathrm{d}\vec{u}\|}{\int \|\mathrm{d}\vec{u}\|}\right)^2 = \left(\frac{\int \|\mathrm{d}\vec{\gamma}\|}{\int \|\mathrm{d}\vec{u}\|}\right)^2 \equiv \left(\frac{L_{\vec{\gamma}}}{L_{\vec{u}}}\right)^2. \quad (12)$$

The inequality results from replacing the maximum by a spatial average. The integration path is the quarter-circle $u_1^2 + u_2^2 = 1$ from $\vec{u} = (1, 0)$ to $\vec{u} = (0, 1)$ and hence $L_{\vec{u}} = \frac{\pi}{2}$. For any curve $\gamma(\vec{u})$ we have the freedom to re-parameterize such that $\lambda(\vec{u})$ is constant over the curve. The above inequality can then be changed into an equality,

$$\min_{\vec{\gamma}} \max_{\vec{u}} T(\vec{u}) = (4/\pi^2)(\min_{\vec{\gamma}} L_{\vec{\gamma}})^2. \quad (13)$$

The problem is reduced to finding a curve $\vec{\gamma}(\vec{u})$ of minimal length with the constraint $\gamma_w(\vec{u}) = \sqrt{(M\vec{h})_w(\vec{u})}$ where $M(t_1, t_2, r)$ is

$$M = \begin{array}{c|ccc} w\backslash\psi & \{0\} & \{1\} & \{0,1\} \\ \hline \emptyset & (1-t_1)(1-r) & (1-t_1)(1-r) & (1-t_2)^2 \\ \{0\} & t_1(1-r) & (1-t_1)r & t_2(1-t_2) \\ \{1\} & (1-t_1)r & t_1(1-r) & t_2(1-t_2) \\ \{0,1\} & t_1 r & t_1 r & t_2^2 \end{array}. \quad (14)$$

4.2 Geodesics

Length-minimizing curves are obtained by solving the geodesic equations for the appropriate metric. In our case the constraint $\gamma_w(\vec{u}) = \sqrt{(M\vec{h})_w(\vec{u})}$ causes complications. If we write $M = [m_1, m_2, m_3]$ then $\vec{g} = M\vec{h}$ is a convex combination of the column vectors m_1, m_2, m_3. The allowed space of $\vec{g}$ is anywhere inside the triangle shown in Fig. 2. We switch from variables (u_1, u_2) to s_1, s_2 with $0 \leq s_1 \leq 1, 0 \leq s_2 \leq 1 - s_1$.

$$\vec{g}(s_1, s_2) \triangleq m_1 + s_1(m_2 - m_1) + s_2(m_3 - m_1). \quad (15)$$

The marking assumption yields $\vec{u} = (1, 0) \Rightarrow \vec{h} = (1, 0, 0)$ and $\vec{u} = (0, 1) \Rightarrow \vec{h} = (0, 1, 0)$. In terms of $\vec{g}(s_1, s_2)$ this means $\vec{g}(1, 0) = m_1$ and $\vec{g}(0, 1) = m_2$. We are looking for the shortest path from the lower left corner (m_1) of the triangle to the lower right corner (m_2). An infinitesimal change in $\mathrm{d}\gamma_w$ is given by

$$\mathrm{d}\gamma_w = \frac{\mathrm{d}g_w}{2\sqrt{g_w}} = \frac{(m_{2,w} - m_{1,w})\mathrm{d}s_1 + (m_{3,w} - m_{1,w})\mathrm{d}s_2}{2\sqrt{g_w}}. \quad (16)$$

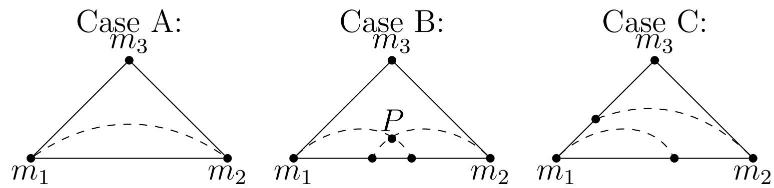


Figure 3: *The three cases we encounter for the way the geodesics intersect.*

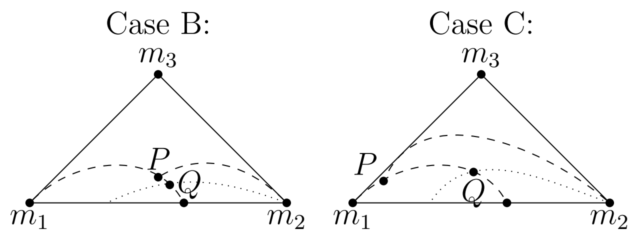


Figure 4: *The optimal path in both cases is m_1−P−m_2 over the dashed lines (geodesics). In case C the geodesic from m_2 is tangent to the left side of the triangle.*

This allows us to define the appropriate metric $G(s_1, s_2)$,

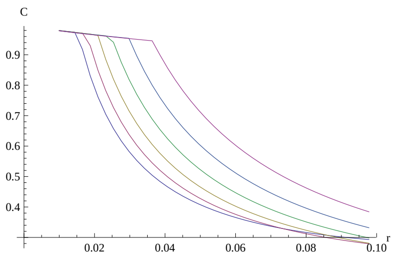
$$\|\mathrm{d}\vec{\gamma}\|^2 = G_{11}(\mathrm{d}s_1)^2 + G_{22}(\mathrm{d}s_2)^2 + 2G_{12}\mathrm{d}s_1\mathrm{d}s_2. \tag{17}$$

See the full paper for details on the geodesic computations. We want to find the shortest path between m_1 and m_2 that is fully inside the triangle. There are three cases. In case A we are done since the direct geodesic is the shortest path. For B and C the optimal paths are shown in Fig. 4. Any geodesic starting from m_2 with a smaller initial slope has to cross the maximum-slope geodesic from m_1 in a point Q. From Q the optimal path to m_1 is to follow the geodesic; but at P you could have done better by going directly from m_2 to P on the geodesic. We use the length of the optimal path to compute the capacity,

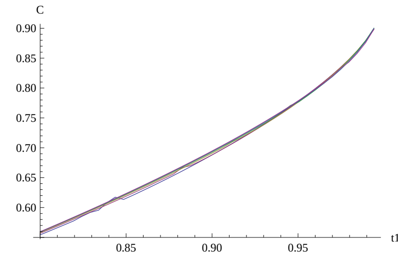
$$C_{2,\infty}^{\mathrm{CDM}} = \frac{1}{2c^2 \ln 2} \frac{4}{\pi^2} L_{\mathrm{opt}}^2. \tag{18}$$

4.3 Results

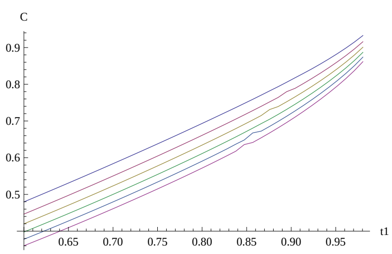
Fig. 5 shows the ratio $C = C_{2,\infty}^{\mathrm{CDM}}/C_{2,\infty}^{\mathrm{RDM}}$ between the asymptotic capacities for the CDM and the RDM as a function of t_1, t_2, r. It turns out that the asymptotic capacity depends on the three attack parameters in a nontrivial way. Obviously, the capacity is an increasing function of t_1 and t_2, and a decreasing function of r. For r close to zero and t_1 close to 1, the capacity has very weak dependence on t_2. This can be understood from the fact that we are close to the Marking Assumption: when the MA holds, all the attack models for $q = 2$ are equivalent. In Fig. 5a we see a transition from linear behavior as a function of r (with almost total insensitivity to t_2) to nonlinear behavior (with dependence on t_2). The transition point depends on t_2.



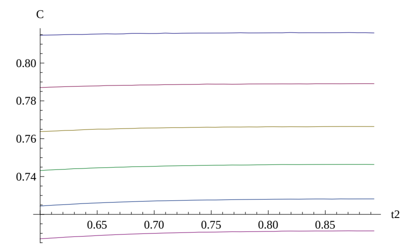
(a) C *vs* r *for fixed* $t_1 = 0.999$ *for* $t_2 = \{0.8, 0.82, 0.84, 0.86, 0.88, 0.9\}$



(b) C *vs* t_1 *for fixed* $r = 0.01$ *for* $t_2 = \{0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8\}$



(c) C *vs* t_1 *for fixed* $t_2 = 0.5$ *for* $r = \{0.01, 0.02, 0.03, 0.04, 0.05\}$



(d) C *vs* t_2 *for fixed* $t_1 = 0.9$ *for* $r = \{0.01, 0.02, 0.03, 0.04, 0.05\}$

Figure 5: *The ratio* $C = C^{\rm CDM}_{2,\infty}/C^{\rm RDM}_{2,\infty}$ *for* $q = 2$.

5 Discussion

We have investigated the asymptotic channel capacity in the CDM. For general alphabet size q it turns out to be very difficult to compute this quantity. We have shown how the previously obtained capacity for the RDM [4] follows as a limiting case of the CDM. For the binary alphabet we have shown how the problem of computing the channel capacity reduces to finding a constrained geodesic between two points. We have presented numerical solutions to this problem. The asymptotic capacity depends on the three attack parameters t_1, t_2, r in a nontrivial way. The graphs show a regime close to the Marking Assumption, in which the $C^{\rm CDM}_{2,\infty}$ is practically independent of t_2.

References

[1] E. Amiri and G. Tardos. High rate fingerprinting codes and the fingerprinting capacity. In *SODA 2009*, pages 336–345.

[2] N.P. Anthapadmanabhan, A. Barg, and I. Dumer. Fingerprinting capacity under the marking assumption. *IEEE Transaction on Information Theory – Special Issue on Information-theoretic Security*, 54(6):2678–2689.

[3] O. Blayer and T. Tassa. Improved versions of Tardos' fingerprinting scheme. *Designs, Codes and Cryptography*, 48(1):79–103, 2008.

[4] D. Boesten and B. Škorić. Asymptotic fingerprinting capacity for non-binary alphabets. In *Information Hiding 2011*, volume 6958 of *LNCS*, pages 1–13. Springer, 2011.

[5] A. Charpentier, F. Xie, C. Fontaine, and T. Furon. Expectation maximization decoding of Tardos probabilistic fingerprinting code. In *Media Forensics and Security*, volume 7254 of *SPIE Proceedings*, page 72540, 2009.

[6] Y.W. Huang and P. Moulin. Saddle-point solution of the fingerprinting capacity game under the marking assumption. In *ISIT 2009*.

[7] Y.W. Huang and P. Moulin. Saddle-point solution of the fingerprinting capacity game under the marking assumption. In *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2009.

[8] M. Kuribayashi. Tardos's Fingerprinting Code over AWGN Channel. In R. Böhme, P.W.L. Fong, and R. Safavi-Naini, editors, *Information Hiding*, volume 6387 of *LNCS*, pages 103–117. Springer, 2010.

[9] M. Kuribayashi. A new soft decision tracing algorithm for binary fingerprinting codes. In T. Iwata and M. Nishigaki, editors, *IWSEC*, volume 7038 of *LNCS*, pages 1–15. Springer, 2011.

[10] M. Kuribayashi, N. Akashi, and M. Morii. On the systematic generation of Tardos's fingerprinting codes. In *MMSP 2008*, pages 748–753.

[11] T. Laarhoven and B.M.M. de Weger. Optimal symmetric Tardos traitor tracing schemes. 2011. `http://arxiv.org/abs/1107.3441`.

[12] P. Meerwald and T. Furon. Towards joint Tardos decoding: the 'Don Quixote' algorithm. In *Information Hiding*, volume 6958 of *LNCS*, pages 28–42. Springer, 2011.

[13] P. Moulin. Universal fingerprinting: Capacity and random-coding exponents. In *Preprint arXiv:0801.3837v2, avilable at `http://arxiv.org/abs/0801.3837`*, 2008.

[14] K. Nuida, S. Fujitsu, M. Hagiwara, T. Kitagawa, H. Watanabe, K. Ogawa, and H. Imai. An improvement of discrete Tardos fingerprinting codes. *Des. Codes Cryptography*, 52(3):339–362, 2009.

[15] K. Nuida, M. Hagiwara, H. Watanabe, and H. Imai. Optimal probabilistic fingerprinting codes using optimal finite random variables related to numerical quadrature. *CoRR*, abs/cs/0610036, 2006.

[16] G. Tardos. Optimal probabilistic fingerprint codes. In *STOC 2003*, pages 116–125.

[17] B. Škorić, S. Katzenbeisser, and M.U. Celik. Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes. *Designs, Codes and Cryptography*, 46(2):137–166, 2008.

[18] B. Škorić, S. Katzenbeisser, H.G. Schaathun, and M.U. Celik. Tardos Fingerprinting Codes in the Combined Digit Model. *IEEE Transactions on Information Forensics and Security*, 6(3):906–919, 2011.

[19] B. Škorić, T.U. Vladimirova, M.U. Celik, and J.C. Talstra. Tardos fingerprinting is better than we thought. *IEEE Trans. on Inf. Theory*, 54(8):3663–3676, 2008.

[20] F. Xie, T. Furon, and C. Fontaine. On-off keying modulation and Tardos fingerprinting. In *MM&Sec 2008*, pages 101–106.