

# Point Compression for Certificates of Authenticity

Darko Kirovski

Microsoft Research, One Microsoft Way, Redmond, WA 98052

Certificates of authenticity (COAs) are digitally signed physical objects that have a random unique structure which satisfies three requirements: (i) the cost of creating and signing original COAs is small (ii) the cost of exact or near-exact replication of COA's physical structure is several orders of magnitude larger than creating an original, and (iii) the cost of verifying the authenticity of a signed COA is small. COAs were first introduced for arms control verification purposes in the 1970s [1]. Bauder was the first to propose COAs created as a collection of fibers randomly positioned in an object using a transparent gluing material which permanently fixes fibers' positioning [1]. Readout of the random structure of a fiber-based COA is performed using the following fact: if one end-point of a fiber is exposed to light, the other one will illuminate.

During certification, the positions of COA's fibers are digitized and compressed as a message  $f$ . Next,  $f$  is combined with an associated text  $t_0$  by hashing  $t_0$  using a cryptographic hash - we denote this hash as  $t$ , and then encrypting  $f$  using  $t$  as a key. The resulting message  $m$  is then signed using issuer's private RSA-key, and finally, the resulting RSA signature  $s$  is imprinted on the COA as a barcode. Each COA instance is associated with an object whose authenticity the issuer wants to vouch. COA verification involves the following tasks. The verifier initially scans  $t_0$  and hashes it to create  $t$ . It also scans the barcode and decodes  $s$ . Next, the verifier performs the RSA signature verification on  $s$  using issuer's public RSA-key and obtains  $m$ . The verifier computes  $f$  by decrypting  $m$  using  $t$  as a key. Finally, the verifier scans COA's statistical properties, creates their presentation  $f'$ , and compares  $f'$  to the extracted  $f$ . If their similarity surpasses a certain threshold, the verifier declares an authentic certificate and vice versa.

In order to counterfeit protected objects, the adversary needs to either: obtain the private key of the issuer, or devise a manufacturing process that can exactly replicate an already signed COA instance, or misappropriate signed COA instances. From that perspective, COA can be used to protect objects whose value roughly does not exceed the cost of forging a single COA instance.

Since barcode capacity is limited, the goal of any COA system is to contain in  $m$  as much information about the random structure of the physical object as possible. To address this issue, we have developed a point compression algorithm that consists of several phases. First, given that a particular region  $S_i$  of COA  $S$  is illuminated, we compute the pdf that a particular point in  $S - S_i$  is illuminated. The COA model is constant for all instances and is, hence, hardwired into the verifier. Next, we derive an arithmetic coder which uses the obtained pdf in order to encode near-optimally a vector between two points. Finally, we model the optimization problem of compressing the positions of as many as possible illuminated points using a fixed  $|m|$ . For each illuminated point  $u \in S - S_i$ , we create a node  $n_u$ . A directed edge  $e(u, v)$  from node  $n_u$  to node  $n_v$  is weighted with the optimal bit-length of the codeword that encodes the vector that points to  $v$ ,  $\omega(e(u, v))$ , conditioned on the fact that  $u$  is already encoded. We denote this graph as  $G(N, E, \Omega)$ , where  $N$ ,  $E$ , and  $\Omega$  represent the set of nodes, directed edges, and their weights. We search for a maximum sized subset of nodes  $N^* \subset N$  with a permutation  $n_{\pi(1)}^*, \dots, n_{\pi(l)}^*$  such that the sum of weights along this path is smaller than  $|m|$ . This problem is NP-complete. We developed an efficient least-constraining most-constrained iterative heuristic that aims at solving this problem with a fixed run-time of one second. The compression ratios obtained using our algorithm were 15-25% lower than the ratios obtained using straightforward point-compression algorithms [2].

[1] D.W. Bauder. Personal Communication.

[2] D. Kirovski. Towards Automated Detection of COAs. Microsoft Research, technical report, 2003.