

BIRN ID generation and Link Table Maintenance

As a step of patient de-identification process , as required by HIPPA, the patient IDs assigned by individual labs should not be used as identifiers. To achieve this, a new set of random identifiers independent of patient IDs (BIRN IDs) needs to be generated. The linkage between one-way hash of the patient ID and BIRN ID is kept in a link table, which is only accessible via the institutions' security officer. The standard pseudo random number generators create repeating sequence of pseudo random number sequences depending on the random seed used. The most common random number generator seeds are time stamps , MAC addresses and combination of them. Guessing the random seed to create the same sequence of random numbers is quite feasible and there are known security breaches of this sort. Since most asymmetric key generation algorithms (as used in SSL, HTTPS, for example) need a pseudo random number generator (PRNG) which produces non-deterministic output and the seed is unpredictable and the output of the secure PRNG is cryptographically strong sequences, there are PRNG algorithms developed by security software companies.

One such algorithm is SHA1PRNG as provided by Sun in Java Development Kit, which follows IEEE P1363 standard, Appendix G.7: "Expansion of source bits", and uses SHA-1 as the foundation of the PRNG. It computes the SHA-1 hash over a true-random seed value concatenated with a 64-bit counter which is incremented by 1 for each operation. From the 160-bit SHA-1 output, only 64 bits are used. SHA-1 the Secure Hash Algorithm, as defined in Secure Hash Standard, NIST FIPS 180-1. SHA1PRNG is cryptographically strong pseudo-random number minimally complies with the statistical random number generator tests specified in FIPS 140-2, Security Requirements for Cryptographic Modules, section 4.9.1.

There is a reference implementation for BIRN ID , patient ID one-way hash (MD5) generation, and link table maintenance program written in Java available under CVS (cvs checkout birn-id_gen).

It is guaranteed that the BIRN ID created is not repeated within the institution. The secure random number generated for BIRN ID has 8 digits, produces non-deterministic output and the seed is unpredictable and the output of the secure random number generator is cryptographically strong sequences (RFC 1750: Randomness Recommendations for Security). Retrieval of the patient (clinical) IDs from the MD5 one-way patient ID hashes is cryptographically infeasible (ie. they cannot be reversed).