Use of Hamming Weights Instead of Uniform Distributions to Analyze a Set of Strings for Randomness

Abstract *- There are a number of reasons why one may desire to have a random string, including, but not limited to: sources of entropy in a system, good cryptographic codes, statistical modeling of a random distribution. Historically when studying if a bit string is random, one would model that string as a uniform distribution with equal probability of each bit being one or zero. While this method works well for a single bit string, we have identified that for an ensemble of bit strings produced by a single generator function, the Hamming Weight (HW) of the bit strings can be a more effective statistic. Using properties of a bit string, with each bit position being limited to a one or a zero, for any bit string of* N*-bits its HW will be restricted to the range* $[0, N]$*. Likewise a histogram of its Hamming Weights binned from HW=*0 *to HW=*N *will follow a binomial distribution with probability* p*=½. In this work we present a method for reasoning about ensembles of bit strings for randomness using the Hamming Weight histogram of the bit strings.*

Keywords: Hamming weight, bit string, randomness, big data statistic, security, computer security

AUTHORS

1:Rendon, Joshua-Southern Methodist University, Darwin Deason Institute for Cyber Security 2:Thornton, Mitchell-Southern Methodist University, Darwin Deason Institute for Cyber Security 3:Thornton, Micah-Southern Methodist University, Darwin Deason Institute for Cyber Security 4:Pham, Gavin-Southern Methodist University, Darwin Deason Institute for Cyber Security