**Lecture 4 Notes**

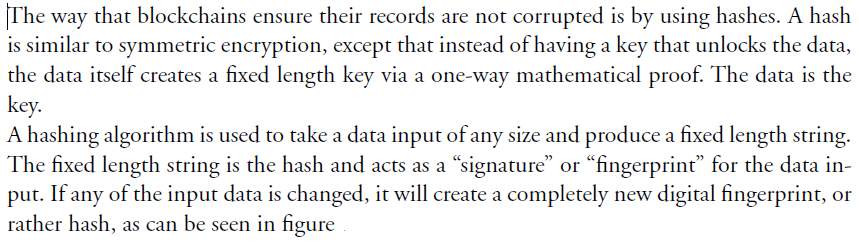
Slide 4

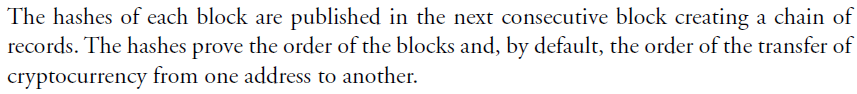
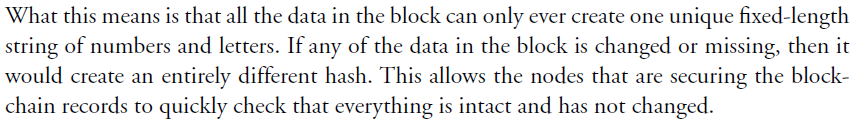
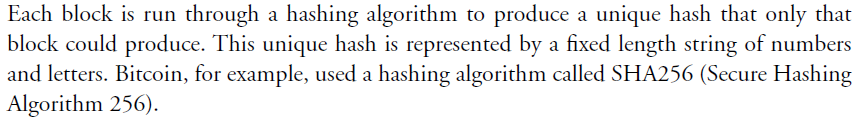
* Hans Peter Luhn, an engineer at IBM coined Hashes in 1950’s as a way to organize numbers and text.

Slide 5

* A cryptographic hash function takes data and essentially translates it into a string of letters and numbers.
* Data goes into a hash function, the function runs, and a string of letters and numbers is produced, that string is called a hash.
* In the Bitcoin blockchain hashes are 256 bits, or 64 characters.
* Input can be extremely short or almost infinitely long and you will still get a unique output string of uniform length.

Slide 6





Slide 7

Cryptographic hashing is a method for transforming large amounts of data into short numbers that are di cult to imitate. One can convert a text or a picture, which represents a variable-length bit sequence, to produce a fixed-length bit sequence in the form of a hash. Hashes are mostly used in combination with digital signatures. These functions ensure data integrity.

Hashing in the Bitcoin Network is used for four processes:

1. encoding wallet addresses;
2. encoding transactions between wallets;
3. verifying and validating the account balances of wallets; and for the consensus mechanism
4. “Proof-of-Work.”

The Bitcoin Network uses SHA (Secure Hash Algorithm), such as SHA-256. An important property of hashes is that if one single bit of input data is changed, the output changes significantly, which makes it easy to detect small changes in large text files, for example. As you can see from the example below, an entirely different hash gets generated when we change only one letter. This is based on the so-called avalanche effect, and it is useful for easily providing data integrity. An entirely different string results from hashing the hash. “Avalanche effect” describes the behaviour of a mathematical function where even a slight change in an input string should cause the resulting hash value to change drastically. This means that in a document of several hundred pages, if one ads only one word, or even a comma, the whole hash will change. A document‘s hash value can, therefore, serve as a cryptographic equivalent of the document – a digital fingerprint. This is why one-way hash functions are central to public-key cryptography. When producing a digital signature for a document, we no longer need to encrypt the entire document with a sender‘s private key, which can take a lot of time. It is sufficient to compute the document‘s hash value instead.

Slide 10

### **Message Digest (MD)**

MD5 was most popular and widely used hash function for quite some years.

* The MD family comprises of hash functions MD2, MD4, MD5 and MD6. It was adopted as Internet Standard RFC 1321. It is a 128-bit hash function.
* MD5 digests have been widely used in the software world to provide assurance about integrity of transferred file. For example, file servers often provide a pre-computed MD5 checksum for the files, so that a user can compare the checksum of the downloaded file to it.
* In 2004, collisions were found in MD5. An analytical attack was reported to be successful only in an hour by using computer cluster. This collision attack resulted in compromised MD5 and hence it is no longer recommended for use.

### **Secure Hash Function (SHA)**

Family of SHA comprise of four SHA algorithms; SHA-0, SHA-1, SHA-2, and SHA-3. Though from same family, there are structurally different.

* The original version is SHA-0, a 160-bit hash function, was published by the National Institute of Standards and Technology (NIST) in 1993. It had few weaknesses and did not become very popular. Later in 1995, SHA-1 was designed to correct alleged weaknesses of SHA-0.
* SHA-1 is the most widely used of the existing SHA hash functions. It is employed in several widely used applications and protocols including Secure Socket Layer (SSL) security.
* In 2005, a method was found for uncovering collisions for SHA-1 within practical time frame making long-term employability of SHA-1 doubtful.
* SHA-2 family has four further SHA variants, SHA-224, SHA-256, SHA-384, and SHA-512 depending up on number of bits in their hash value. No successful attacks have yet been reported on SHA-2 hash function.
* Though SHA-2 is a strong hash function. Though significantly different, its basic design is still follows design of SHA-1. Hence, NIST called for new competitive hash function designs.
* In October 2012, the NIST chose the Keccak algorithm as the new SHA-3 standard. Keccak offers many benefits, such as efficient performance and good resistance for attacks.

### **RIPEMD**

The RIPEMD is an acronym for RACE Integrity Primitives Evaluation Message Digest. This set of hash functions was designed by open research community and generally known as a family of European hash functions.

* The set includes RIPEMD, RIPEMD-128, and RIPEMD-160. There also exist 256, and 320-bit versions of this algorithm.
* Original RIPEMD (128 bit) is based upon the design principles used in MD4 and found to provide questionable security. RIPEMD 128-bit version came as a quick fix replacement to overcome vulnerabilities on the original RIPEMD.
* RIPEMD-160 is an improved version and the most widely used version in the family. The 256 and 320-bit versions reduce the chance of accidental collision, but do not have higher levels of security as compared to RIPEMD-128 and RIPEMD-160 respectively.

### **Whirlpool**

This is a 512-bit hash function.

* It is derived from the modified version of Advanced Encryption Standard (AES). One of the designer was Vincent Rijmen, a co-creator of the AES.
* Three versions of Whirlpool have been released; namely WHIRLPOOL-0, WHIRLPOOL-T, and WHIRLPOOL.

Slide 11

* MD 5: It produces a 128-bit hash. Collision resistance was broken after ~2^21 hashes.
* SHA 1: Produces a 160-bit hash. Collision resistance broke after ~2^61 hashes.
* SHA 256: Produces a 256-bit hash. This is currently being used by [bitcoin](https://blockgeeks.com/guides/what-is-bitcoin/).
* Keccak-256: Produces a 256-bit hash and is currently used by [ethereum](https://blockgeeks.com/guides/ethereum/).

Slide 12

* SHA256 is used in Bitcoin mining – to construct the Bitcoin blockchain.
* Secure Hash Algorithm (SHA) that generates 256 bit message digest.
* A part of SHA-2, a set of cryptographic hash functions designed by United States National Security Agency (NSA).