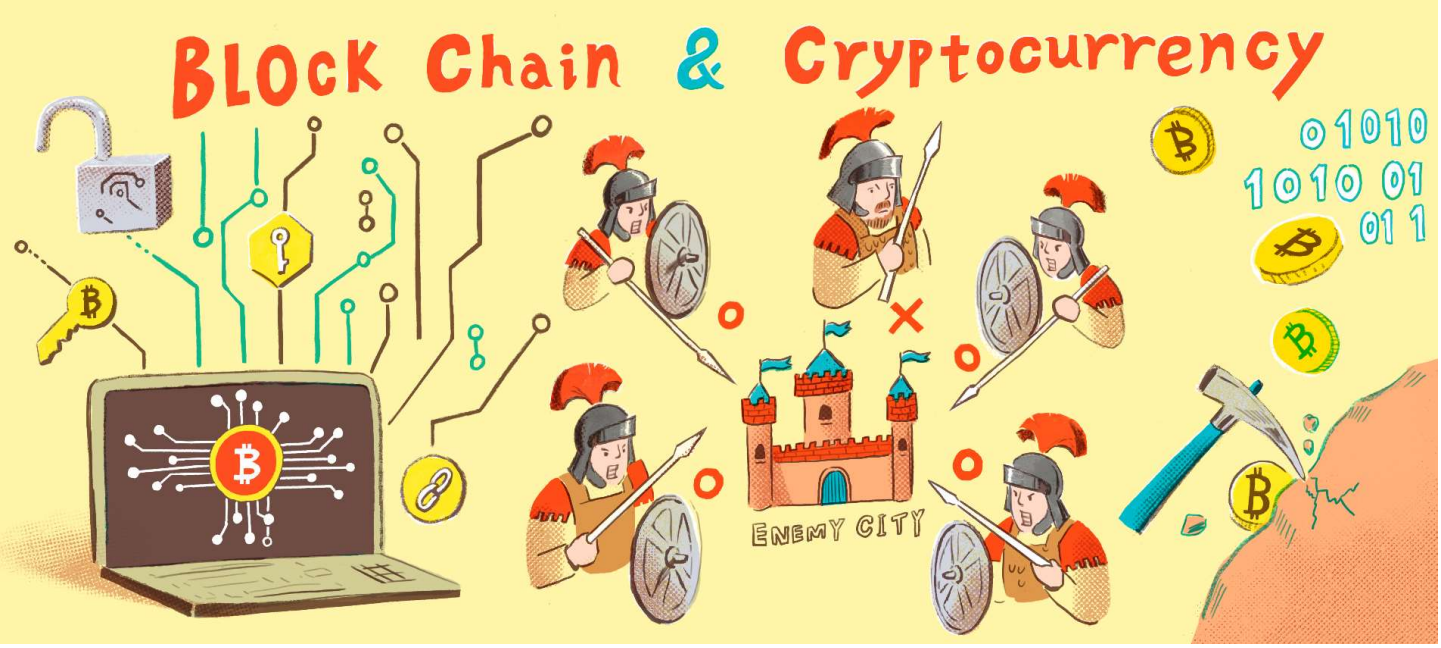
블록체인

블록체인과 암호화폐

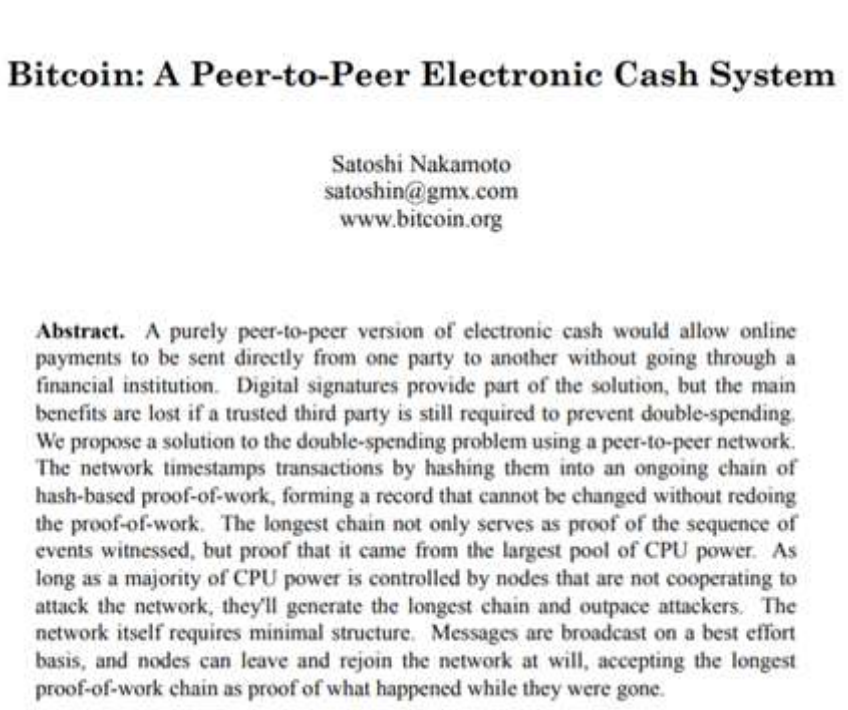
2021년 10월 26일

김종락



들어가며

2008년 10월 31일 사토시 나카모토Satoshi Nakamoto라는 필명의 저자가 “Bitcoin: A Peer-to-Peer Electronic Cash System”이라는 제목의 9쪽 짜리 논문([그림1])을 암호 메일링 리스트에 올렸다. 논문 초록은 금융 당국을 거치지 않고도 온라인 결제가 가능한 전자화폐가 가능해질 것이라는 내용으로 시작한다. 이는 디지털 서명을 통해서도 일부 가능하지만, 여전히 제 3자가 개입할 수 있다는 단점이 있다. 이를 해결하기 위한 가장 큰 문제는 이중사용double-spending의 문제이다. 현재는 우리가 사용하는 금융 거래 기록을 은행에서 관리하기 때문에 이미 사용한 금액을 재사용할 수가 없다. 그런데, 만일 은행이 없다면 이런 관리를 어떻게 해결해야 할까?



Bitcoin: A Peer-to-Peer Electronic Cash System

Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

그림1 사토시의 비트고인 논문

여기서부터 많은 문제가 발생하고 이를 해결하기 위한 방법으로 사토시는 해쉬 기반의 작업증명Proof-of-Work을 제안했다. 각각의 거래 내용을 기다란 체인에 계속 붙임으로써 제 3자가 변경할 수 없도록 하는 방법이다. 블록체인의 전반적인 개념을 좀 더 자세히 설명해 보자.

블록체인의 개념

블록체인의 정의는 다양하다. 영문 위키피디아에서는 블록체인을, 암호를 이용해 생성된 블록이라 불리는 거대한 기록물records들의 모임이라고 정의하고 있다. 좀 더 부연하면, 노드라고 불리는 참가자들로 이루어진 네트워크상에서 신뢰가 없더라도 완전히 분산된 P2PPeer to Peer 기반의 변경 불가능한 데이터 스토리지로 블록체인을 정의하기도 한다.

따라서 블록체인의 특징은 크게 4가지로 나눈다. 자율 분산 시스템, 리더가 없는 구조, 지갑 주소, 트랜잭션(거래)이 그것이다. 각각을 살펴보면 첫째 모든 노드가 같은 데이터를 복사하여 작동하므로 일부 노드가 고장 나더라도 전체 시스템의 동작에 지장을 주지 않는다. 둘째 특정 노드가 리더 역할을 하지 않는 민주적인 구조이다. 셋째와 넷째, 공개 키와 비밀 키로 이루어진 지갑주소가 있어 각각의 거래 내역이 블록체인에 기록된다. 해시(암호화)와 공개키 암호의 특징 덕분에 블록체인은 위변조 없이 믿을 만한 거래가 이루어지고 있다고 할 수 있다. 네 가지 특징에 더해 짧은 문자열을 보낼 수 있다는 특징이 추가된 대표적인 블록체인인 '스마트 계약'이 이더리움이다.

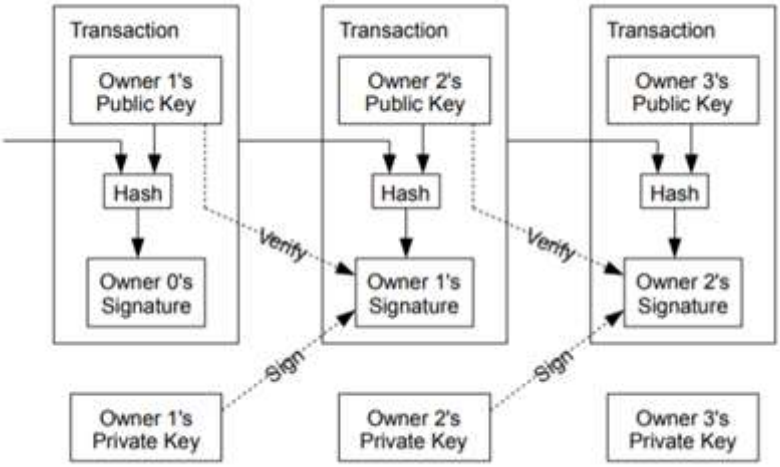


그림2 각 사용자의 거래 장부 및 전자 서명

[그림2]는 각 사용자의 트랜잭션(거래) 및 전자 서명signature을 기술한 것이다. 각 거래에는 금액에 대한 정보와 서명이 들어있다. 그림에서 보듯이 사용자 1은 자신의 거래 내용을 사용자 2에게 전달할 때 거래장 마지막에 사용자 1의 서명을 넣어 이 거래가 자신이 거래한 것임을 증명하였다. 제 3자에 대한 왜곡을 피하기 위해 사용자 1의 거래와 사용자 2의 공개키를 해시(암호화)한 것에 대하 사용자 1이 서명한 것이다. 이것을 체인으로 연결함으로써 중간에 제 3자가 개입한다고 하더라도 모든 해시값을 복호화하는 것은 거의 불가능해진다.

거래를 안전하게 진행하기 위해서는 이전 사용자들이 이미 사용한 내용을 현재 사용자가 다시 사용하지 못하도록 해야 한다. 이를 해결하기 위해 사토시는 타임스탬프 서버를 제안했다. 타임스탬프 서버는 아이템의 블록을 해시한 후 시간을 부여하고 이를 공개적으로 포스팅한다. 타임스탬프는 그 시간대에 데이터가 존재했다는 것을 증명한다. 타임스탬프 기술은 분산 타임스탬프 기술의 하나이다. 개별 거래에 포함된 타임스탬프의 정밀도는 낮을 수 있으나, 다수의 거래의 타임스탬프를 블록 단위로 비교하여 비슷한 시각의 타임스탬프를 선택하면, 신뢰할 수 있는 정밀도와 정확성을 보장할 수 있다.

따라서 블록은 크게 블록 헤더와 트랜잭션 부분으로 나뉜다. 블록 헤더는 6개의 정보, 즉 블록의 버전 정보, 이전 블록 헤더의 해시값, 해당 블록에 포함된 트랜잭션의 머클 루트(이진 트리 구조) 해시값, 블록이 생성되었을 때의 타임스탬프, 채굴의 난이도, 그리고 채굴 시 주어진 논스 등으로 이루어져 있다. 이 전체의 블록이 SHA2-256 등의 해시 함수에 의하여 해쉬된 후 다음 블록에 연결된다. 따라서 블록체인은 각 단계별로 해시를 취함으로써 이것을 역으로 알아내는 것이 불가능하도록 하는 분산 암호 기술이라고 할 수 있다.

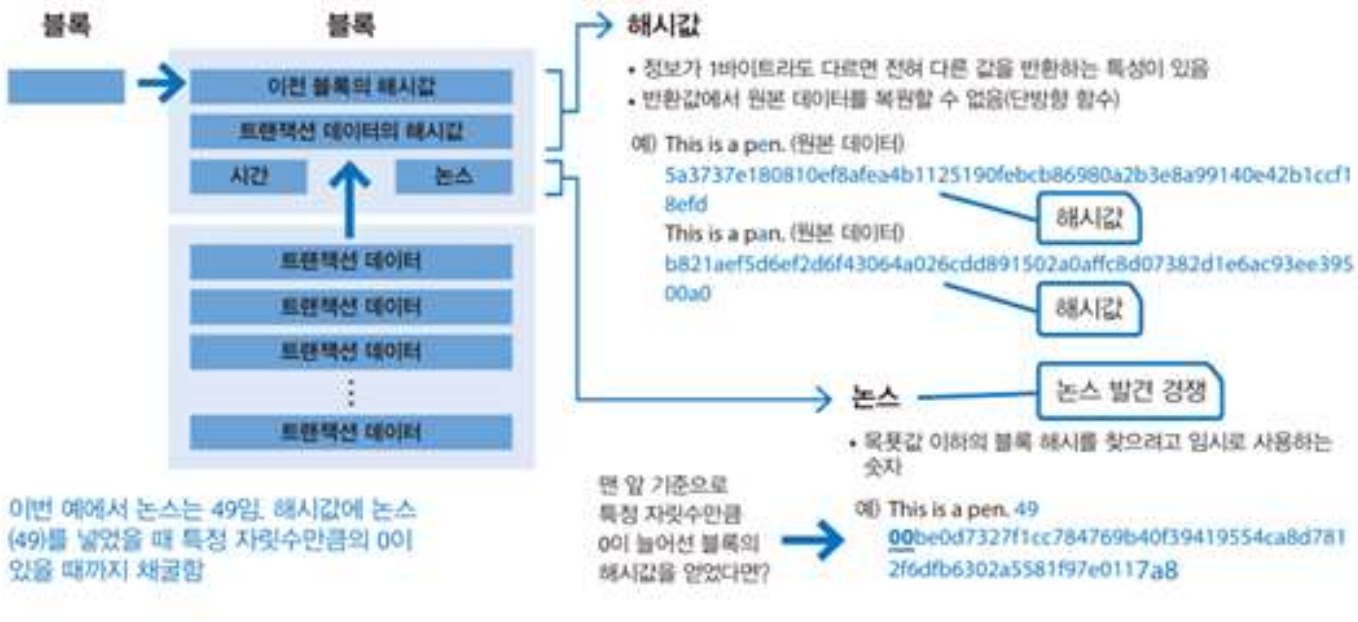


그림3 작업 증명의 기본 개념

스기이 야스노리, 『한 권으로 끝내는 블록체인 교과서』, 제이펍, 2020, 128쪽

작업 증명Proof of Work, POW은 [그림3]처럼, 제시된 목표값보다 작거나 같은 해시값을 찾기 위해 무수히 반복함으로써 해당 작업에 참여했음을 증명하는 합의 알고리즘이다. 작업 증명은 어떤 트랜잭션이 발생했을 때 해당 트랜잭션이 유효한지 판단하고 새로운 블록이 진짜인지 혹은 가짜인지 검증하는 역할을 한다. 한편, 언론에서 많이 언급되는 채굴Mining은 임의의 값을 대입하여 얻은 결과 값이 논스Nonce의 해시값으로 제시된 타겟보다 작은 결과 값이 나올 때까지 무한 반복하는 작업이다. 해시함수는 입력값이 약간만 달라도 결과가 전혀 다르게 나오기 때문에, 일반적으로 제시된 타겟보다 작은 결과 값이 나오도록 하려면 끊임없이 값을 대입해야 한다. 혹자는 이를 가리켜 어려운 수학문제를 푸는 것이라고 말하지만, 사실은 문제가 풀릴 때까지 값을 무한히 반복하는 것에 가깝다. 따라서 지구의 여러 곳에서는 값싼 전기값을 이용하여 수십 대의 고성능 컴퓨터를 돌려서 채굴하고 이에 따른 보상으로 코인을 획득하는 사람들이 꽤 있다.

블록체인과 암호 함수

블록체인에서 해시 함수의 역할이 크기 때문에 좀 더 자세히 살펴보자. 해시 함수는 임의의 입력값에 대하여 동일한 크기의 값을 출력값으로 배출하는데, 주어진 출력값으로는 원래의 입력값을 찾기 어렵게 만드는 함수이다. 해시 함수는 서명을 하는데 쓰이는 암호 함수이다. 대표적인 해시 함수는 SHASecure Hash Algorithm 시리즈의 SHA2-256이며 출력값은 256비트이다. 실제로는 16진수로 표현하기 때문에 결과값은 64자리로 나온다. 예를 들어 다음 사이트에 들어가서

홍길동이라고 치면,

9c6b9b1b1627f3120e0730c6d2cfa71040fd03747bde2755e8b5e4dbf2bee262

라고 해시값이 나온다. 아주 일부가 변경된 홍길도라고 치면,

09324aed12ccaa8867621990e6b255e314f2854693576a0d6c192e05fe7edaa7

라는 전혀 다른 모양의 해시값이 나온다. 이처럼 좋은 해시 함수는 입력값의 차이가 작더라도 출력값은 전혀 다르게 나온다. 즉 입력값을 평서문이라고 하고 출력값을 암호문이라고 할 때, 다수의 "평서문, 암호문" 쌍을 알게 되더라도 새로운 암호문에 대응하는 평서문을 예측하기 어려워야 좋은 암호라고 할 수 있다. 현재는 SHA2보다 보안이 강화된 3세대 해시 함수인 SHA3가 사용되고 있다.

블록체인에 적용된 공개 키 암호는 타원 곡선 암호Elliptic Curve Cryptography를 자주 사용한다. 공개키 암호는 개인키private key와 공개키public key로 이루어져 있다. 개인키는 자신만이 아는 암호이고 공개키는 전화번호부 책처럼 공개된 것이다. 1970년대 RSA 암호에서 시작된 공개키 암호는 지금까지 사용되고 있다. 예를 들어 공개키를 이용하여 갑이 을에게 메시지 m을 보낸다고 하자. 그러면 갑은 을의 공개키를 m과 결합하여 을에게 보낸다. 그러면 을은 자신의 개인키를 이용하여 갑이 보낸 메시지 m을 복원하는 방식이다. 현재까지 알려진 공개키 암호는 RSA외에도 타원 곡선 암호, 부호 기반 암호 등이 있다.

Symmetric Key Size (bits)	RSA and Diffie-Hellman Key Size (bits)	Elliptic Curve Key Size (bits)
80	1024	160
112	2048	224
128	3072	256
192	7680	384
256	15360	521

그림4 RSA 암호와 타원 곡선 암호의 공개키 길이

비트코인에서 사용된 공개키 암호는 타원 곡선 암호다. 타원 곡선 암호는 타원 곡선의 식인 $y^2 = x^3 + ax + b$를 만족하는 (x, y)점들의 집합이다. 이 점들에 대하여 대수적인 덧셈이 정의되어 있다. 이 연산에 관한 이산로그 문제[1] Discrete Log Problem에 다항식 알고리즘이 아직 발견되지 않아 암호에 쓰이고 있다.

[1] 숫자 g가 주어지고 $y = g^x (mod p)$ 값을 알 때, x를 구하는 것을 이산로그 문제라고 한다.

[2] 여기서는 p가 소수이므로 덧셈과 곱셈에 대한 연산을 $mod p$ 연산(p로 나눈 나머지 값)으로 진행하면 된다.

또한 타원 곡선 암호는 RSA 암호보다 키 길이가 짧다는 장점이 있다. [그림4]처럼, 128 비트의 안전성(즉 2^128의 계산이 필요함)을 가정할 때, RSA 암호는 3072의 비트의 키가 필요한 반면 타원 곡선 암호는 256 비트면 충분하다. 따라서 거래가 빈번한 전자화폐에서 타원 곡선 암호를 사용하는 것은 자연스럽다.

비트코인에서 사용된 타원 곡선의 식은 실수 위에서 정의되며 $y^2 = x^3 + 7$로 표현된다. 실제로는 유한체인 Z_p(단, $p = 2^256 - 2^32 - 2^9 - 2^8 - 2^7 - 2^4 - 1$인 소수) 위의 수[2]로서 x, y가 $y^2 = x^3 + 7$을 만족해야 한다. 이 타원 곡선은 secp256k1([그림5])이라고 이름이 붙여졌는데 비트코인 이전에는 거의 주목을 받지 못하였다.

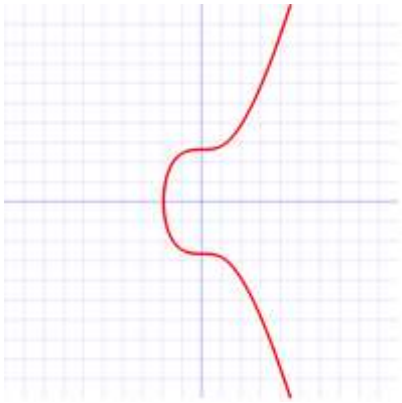


그림5 secp256kI의 실수 위에서의 그래프

블록체인은 퍼블릭 체인과 프라이빗 체인으로 구분된다. 퍼블릭 체인은 참여를 원하는 는 노드들은 네트워크에 참여가 가능하므로 노드 수를 정확히 알 수 없다. 프라이빗 체인은 참여를 원하는 노드에 대한 승인 과정을 통해 네트워크 참여를 제한하므로 노드 수를 알 수 있다. 프라이빗 체인은 노드 수가 적기 때문에 전체 동작도 빠르다. 프라이빗 체인은 다수결 방식의 합의 알고리즘을 이용하기 때문에 노드 운영자에게 보상이나 트랜잭션 수수료를 줄 필요가 없다. 우리가 잘 아는 비트코인, 이더리움 등은 퍼블릭 체인이며, 하이퍼레저 패브릭Hyperledger Fabric은 프라이빗 체인이다. 프라이빗 체인은 어느 정도 중앙화되기 때문에 탈중앙집권을 추구하는 블록체인과 거리가 멀다는 비판도 있다.

비트코인은 퍼블릭 체인이므로 다수결에 따라 합의를 형성할 수 없다. 얼마나 많은 사람들이 네트워크에 참여하고 있는지 알 수 없기 때문이다. 이를 해결하기 위하여 경제적 보상이라는 개념이 생겼는데, 경제적 보상이 바로 우리가 알고 있는 암호화폐이다. 퍼블릭 체인이 택한 합의는 이른바 비잔티움 장애 허용Byzantine Fault Tolerance, BFT을 합의 형성에 적용한 것이다. 비잔티움 장애는 비잔티움 장군의 문제에서 기인한다. 비잔티움 장군의 문제란, 장군들이 멀리 떨어져 전령을 통해 통신해야 하는데 일부 장군이 배신자가 되어 거짓 정보를 보낼 때, 얼마나 많은 장군들이 충직스러워야 하며, 어떤 규칙에 따라 교신을 하면 이런 문제를 해결할 수 있는지에 관한 문제다. 블록체인에서는 배신하는 사람에게는 보상이 없고 충직한 사용자에게는 보상을 주는 식으로 일정수의 배신자가 있더라도 합의에 도달하도록 유도한다.

블록체인이 비잔티움 장군 문제를 어떻게 극복하는지 좀 더 구체적으로 살펴보자. 10개의 노드 중 2개의 노드가 나쁜 정보를 전파한다고 하자. 그러면 8개는 착한 정보를 지니고 있을 것이다. 이제 원래 블록체인은 두 가지, 즉 착한 노드가 채굴한 블록과 나쁜 노드가 채굴할 블록으로 분기될 것이다. 8개의 착한 블록 채굴자들의 채굴 연산 속도는 2개의 나쁜 블록 채굴자들보다 훨씬 빠르게 진행되어 착한 블록 채굴자들이 채굴한 쪽은 더 긴 블록체인을 형성할 것이다. 작업증명POW을 사용하는 블록체인은 가장 긴 블록체인을 선택하게 되고 나쁜 노드의 정보는 버리게 된다. 따라서 일정한 수의 나쁜 노드가 있더라도 정보의 왜곡은 되지 않는 것이다. 물론 여기서 주의해야 할 것은 실제로는 착한 노드의 수가 아니라 계산력이 빠른 노드들이 다수를 차지해야 정보의 왜곡이 일어나지 않는다는 것이다.

블록체인과 산업분야

블록체인이 가장 활발하게 진행되고 있는 분야는 금융이다. 예를 들어 해외 은행 중 스페인어권을 중심으로 사업하는 스페인의 대형 은행인 산탄데르Santander은행은 2016년 블록 체인을 활용하여 국제 결제 시범 프로그램을 개시해 실증 실험을 진행했다. 독일의 최대 은행 중 하나인 도이치 은행도 블록체인에 많은 관심을 갖고 있다. 도이치 은행이 밝힌 블록체인의 활용 영역은 다음과 같다. 예를 들어, 식별·분할·추적이 가능한 유가 증권의 발행 및 전송, 유가 증권의 수익 및 배당 등의 자동화, 무역후 처리 과정을 보다 효율적으로 수행하기 위한 결제 및 청산, 스마트 계약을 사용한 금융 파생 상품 관리의 간소화 등이 그것이다.

"

각 나라의 금융기관이 힘을 합해 블록체인을 연구하기 위한 국제적인 모임을 R3 컨소시엄이라고 한다.

2015년에 조직된 R3 컨소시엄에는 현재 100여 개가 넘는 세계은행이 참여하고 있다.

"

일본에서도 거대 은행 및 인터넷 뱅킹 모두 블록체인의 활용에 나서고 있다. 예를 들어, 미쓰비시 도쿄 UFJ 은행, 미치이 스미토모 은행, 미즈호 은행, 스미신 SBI 넷 은행 등이 블록체인을 활용해 다양한 결제 업무를 해결하고 있다. 각 나라의 금융기관이 힘을 합해 블록체인을 연구하기 위한 국제적인 모임을 R3 컨소시엄이라고 한다. R3 컨소시엄은 신기술의 인프라 구축 및 규제에 대한 대응 등을 연구하기 위해 미국 벤처 기업인 R3 CEV사가 주축이 돼 2015년에 조직했다. 현재 100여 개가 넘는 세계은행이 참여하고 있다. 설립은행으로는 골드만 삭스와 JP 모건 등 9개 은행이 참여했다.

공공분야에 대한 블록체인의 활용도 서서히 두각을 나타내고 있다. 특히 에스토니아 정부는 인구 130만 명의 작은 북유럽국가이지만 IT 강국을 정책으로 내걸고 있다. 실제로 에스토니아 정부는 의료데이터에 대하여 블록체인 기술을 적용하고자 한다. 민감한 정보를 블록체인으로 보관하여 제 3자의 훼손이나 공격으로부터 막을 수 있을 것이다. 의료데이터뿐 만 아니라, 전 국민을 대상으로 블록체인 기반의 ID 카드를 발급했다. 납세, 선거, 회사 설립, 은행 수속 등을 이 카드를 통해 진행할 수 있다.

KID에서 2020년 1월 실시한 블록체인 도입실태 및 향후 전망에 대한 의견조사에 의하면 현재 블록체인의 주도국은 미국이지만 향후 5년 내에 중국이 주도할 것이라고 응답했다. 이에 비하여 우리나라의 블록체인 경쟁력은 선진국에 비하여 절반 수준이라고 대답했다. 법적 및 제도적 측면의 규제가 발전을 저해하고 있다고 지적된다. 좀 더 아쉬운 것은 대부분의 기업들이 블록체인 도입을 고려하고 있지 않고 있다. 그럼에도 불구하고 정보의 안정성 보장 및 비대칭성 해소에 큰 효과를 볼 것으로 전망하고 있다.

맺으며

블록체인과 관련해 해결할 문제도 많다. 현재 블록체인의 주어진 수학문제를 풀기 위해 컴퓨터를 끊임없이 사용하면서 상당한 전기 소모를 야기하고 있다. 2019년 기준으로 비트코인 채굴을 위한 연간 전기 사용량은 133TWh(테라와트)인데, 이는 26위 말레이시아나(147 TWh)와 28위 스웨덴(131 TWh)의 사이 값이다. 한국은 527 TWh를 기록하고 있다.

또 하나의 문제는 양자 컴퓨터가 실현되면 타원 곡선 암호가 깨진다고 알려져 있기 때문에, 이를 기반으로 한 대부분의 블록체인이 안전하지 않을 것이라는 점이다. 혹자는 아직 10년 혹은 20년 후의 일이고 그때가 되면 현재 블록체인의 알고리즘을 보완하여 진행하면 된다고 하지만, 이는 잘못된 생각이다. 지금까지 진행해온 블록들을 미리 저장해 놓고 10년 후 양자 컴퓨터가 그 정보를 캐내게 되면 각 개인의 거래 정보를 볼 수 있기 때문이다. 따라서 지금부터라도 양자 컴퓨터에도 안전한 블록체인 개발이 시급하다. 2017년부터 시작된 NIST의 양자 후 암호Post-quantum Cryptography 경연대회는 지금 3라운드에 접어들고 있다. 현재 총 7개의 암호가 선정되었고 조만간 일부만 최종 선정될 것으로 보인다. 그 이후부터는 선정된 암호를 기반으로 한 블록체인 연구가 활발할 것으로 전망된다.

또한 블록체인과 인공지능이 결합한다면 탈중앙화와 탈인간화가 동시에 이루어지는 요지경의 세상이 될 것이다. 이를 준비하기 위하여 국내 학계 간 경계를 허물고 다양한 산업군과 협업하여 두 기술이 공존할 수 있는 새로운 생태계를 구축할 필요가 있다. 이것이 향후 대한민국의 미래 먹거리가 될 것이라고 필자는 확신한다.

참고문헌

1. Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”, 2008.
2. 스기이 야스노리,『한 권으로 끝내는 블록체인 교과서』, 이중민 옮김, 제이펍, 2020.
3. 아카하네 요시하루 등 공저,『블록체인 구조와 이론』, 양현 옮김, 위키북스, 2017.