#### Big Data and Blockchain to Improve Performance of Professional Sports Teams

Yuri Hong and Dea-woo Park\*

Department of Convergence Engineering, Hoseo Graduate School of Venture, Republic of Korea

Professional Sports market has been growing rapidly, generating numerous analytical tools applied to the performance of sports teams and players. Moreover, adding to traditional points system, various data points including performance data from new perspectives, medical sources, social network data, etc. have sprung up. Thus, the biggest challenge of professional sports teams is to estimate the player with the best odds and efficiency in a particular day's match and the odds of winning or losing. By using newly developed analytical tools such as Big Data, AI, and Blockchain, real-time management of player's skill, performance, and odds become available to make more accurate predictions. In this paper, the management system to improve the performance of professional sports teams using Big Data and Blockchain is introduced so the supervisors and the general managers use the system in their decision-making in South Korea and Internationally. The real-time data is updated through the communication protocol, based on the encrypted Blockchain, then is delivered to the interest group.

Keywords: Blockchain; Big Data; Crypto; Sports Game Data; Performance

#### I. INTRODUCTION

Big data analysis in professional baseball began to take off with Billy Beane, the general manager of the Oakland Athletics baseball team, famous enough to be made into a movie called "Moneyball." In the wake of the 4th Industrial Revolution, it is increasing trend in the sports industry that Big Data analytics are more actively used to estimate the odds of the professional sports games.

Also, legitimized sports betting games such as fantasy sports as well as professional sports leagues attract more funds to the sports industry, making differentiated information and data valuable to the interest groups in the industry. This new trend is triggering the introduction and development of new technology including Blockchain, AI (Artificial Intelligence) and ICT (Information Communication Technology) in the industry.

In this paper, the management system to improve the performance of professional teams using Big Data, AI, and Blockchain is introduced so the supervisors and the general managers use the system in their decision-making. Moreover, by using the advanced ICT, real-time data is updated through the communication protocol, based on the encrypted Blockchain, then is delivered to the interest group. The system will be designed to provide differentiated value through security algorithm to the paying customers. In this system design, Hyperledger Fabric blockchain technology is used owing to its advantages such as strict management of accounts and selective consensus algorithm.

In Section II, related works such as studies on Big Data and Blockchain are introduced. Section III explains historical analyses and examples of Big Data analysis and Blockchain technology in professional sports. In Section IV, big data and block chain system for professional sports team is designed and using Hyperledger Fabric Technology.

#### II. RELATED WORKS

A. Big Data

<sup>\*</sup>Corresponding author's e-mail: prof\_pdw@naver.com

As the society becomes complicated, new analytical tools has sprung up to more precisely conduct analyses and make predictions, such as Altimetrics, an emergent research area whereby social media is applied as a source of metrics to evaluate scientific impact (Li *et al.*, 2017). Big Data analysis is one of the newly developed tools.

IDC describes big data as "a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high-velocity capture, discovery, and analysis" (Gantz and Reinsel, 2011).

Big data mainly refers to the analysis, processing and mining procedure on massive, isomerous, and unstructured data set (Zhe *et al.*, 2017). Big-data analysis requires a huge amount of data storage. Computing power is also used to handle the big data. Therefore, this can be a huge burden not only for individuals but also for small and medium enterprises.

Contemporary society is characterized by massive amounts of data. The amount of data is transmitted and integrated at a tremendous rate. In addition, the directions and interactions of the significant amount of data becomes more and more complicated and diverse. Thus, computer scientists have proposed a number of novel machine learning and data mining algorithms, data visualization techniques, data reduction techniques, etc., in order to extract and discover meaningful information hidden in the data (Zhe *et al.*, 2017). Big data analysis is, thus, being utilized in various aspects in the modern society, ranging from consumer goods marketing tools, sports statistics, even to cultural aspects, such as movies. Using big data, the factors that contributed to the commercial success of films are analyzed to predict box-office performance (Hwang *et al.*, 2017).

#### B. Blockchain

A blockchain is the technology that allows all members to keep a ledger containing all transaction data and to update their ledgers to maintain integrity when there is a new transaction (Park & Hyuk, 2017). It is very important to secure the reliability of a transaction under a network where a large amount of information is present and numerous computers are connected. More and more, the reliability of a transaction in a network has become possible due to the spread of the Internet and the development of cryptographic techniques. In addition, there is a growing need for authorities to prove and

approve the reliability of transactions on networks where many members are connected.

Under the blockchain, this reliability problem can be solved. The characteristics of the block chain are that they are arranged in blocks by time, and that the distributed chain ledger structure possesses a ledger in which all the transaction records are updated by everyone participating in the block chain network.

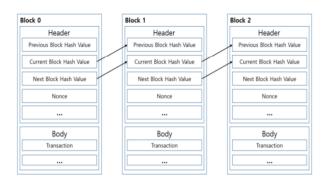


Figure 1. Blockchain Connection Network (Park & Hyuk, 2017)

This way, it is difficult to modulate because it is possible to know when the transaction record is created. Blockchain is a decentralized system, not a central one, so that malicious transactions can be prevented.

The reliability of data is verified using the index method as shown in the figure 1 above (Park & Hyuk, 2017). A block chain creates a block of multiple transactions, and the block contains a hash of the previous block. Each block consists of a header and a body. The header of the block consists of the hash value information of the header of the previous block and the hash value of this transaction. Therefore, if the contents of a certain block are falsified or modulated, the hash value included in the next block changes according to the characteristics of the hash function, and the modulation is disclosed.

Generally, blockchain can be divided into two groups - public blockchain and private blockchain (Yang *et al.*, 2018). When trusted participants organize a consortium, it operates in the form of a private blockchain. In this case, there is no need for a consensus algorithm to ensure reliability such as Proof of Work, which requires time and cost.

#### III. BIG DATA AND BLOCKCHAIN ANALYSIS FOR IMPROVEMENT OF PROFESSIONAL TEAMS' PERFORMANCE

Historically, extensive analysis of professional athlete performance date has been conducted to assist various stakeholders, including professional players, team coaches and managers, recruiting managers, team doctors, nutritionists, and even sports fans, in making their accurate and timely decisions.

### A. Performance Analysis of a Professional Sports Player

The simplest data that historically shows professional athletes and team skills in professional baseball is the two-line Line Score Board as shown in the table I below, showing each team's run totals by inning, and total runs, total hits, and total errors. Later, the Box Score is introduced to more systematically show comprehensive data on the performance of individual players and teams. However, Box Score Board also has limitations in the objective analysis in the cross-sectional and multi-dimensional aspects.

Table 1. Line scores, MLB Blue Jays versus Rays Aug 7, 2010

	1	2	3	4	5	6	7	8	9	R	Н	Е
TB (67-43)	0	0	2	0	3	0	2	0	4	11	9	1
TOR (58-52)	1	2	3	2	5	1	3	0		17	20	2

### B. Big Data Analysis of a Professional Sports Player

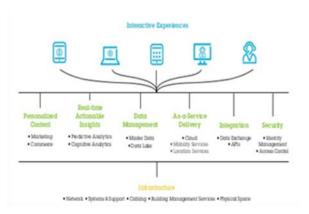
The table II below shows how Big Data analysis is used to analyse a professional sports player using more comprehensive and multi-dimensional data. MLB.COM provides comprehensive data - total 81 items under 12 categories (Game, Month, Timeframe, Opponent, At-Bat, Order, Runners, Position, Inning, Outs, Score, Count) - for each player of all Major League Baseball teams since 1987.

The data in the table II above, being retrieved from the comprehensive data set, shows the performance scores of J. Ramirez, third baseman of Cleveland Indians, against all left-handed pitchers in the games played in 2017. From the data, it is analysed that J. Ramirez is the best hitter of Cleveland team players against left-handed pitchers in that year. The Cleveland coaches can set up strategies centering on Ramirez

if the opponent teams' starter is a left-handed pitcher. Other MLB teams can also consider from this analysis to bring in Ramirez if they have poor track record against left-handed pitchers.

	01.5	0.0			040			70			40			0.5								050
1 Ramirez, J	CLE	3B		1	213	3	34	70	22	1	10	31	15	25	4	3	.329	.3	/1	.582	2	.953
Tampa Bay		AB	R	Н	RBI	88	SO	LOB	AVG	T	oront	0				AB	R	Н	RBI	BB	so	LO
Upton, B, CF		2	1	1	2	1	1	0	.235		Snider,	LF				4	2	2	0	2	0	1
a-Rodriguez, S, PH-3B		2	2	1	0	0	0	0	265	1	Scoba	r, Y, 53				6	1	3	1	0	1	2
Johnson, D., 18		2	2	2	4	2	0	0	273	E	Baurist	J, RF				3	2	1	1	2	1	3
Longoria, 3B		3	0	0	1	0	0	3	285		Wise,	RF				1	0	0	0	0	0	2
b-Kapler, PH-CF		2	0	1	1	0	8	0	207	١	Nels,1	/, CF				4	1	1	1	2	0	3
Joyce, LF		4	0	0	-	1	2	2	210	1	ind, D	Н				5	2	2	3	0	1	5
Aybar, W, DH		4	0	0	1	1	0	2	248	1	10, A,	2B				5	2	3	4	0	1	5
Zobrist, RF		5	0	0	0	0	1	3	266		1-McE	lonald,	lo, PR-	28		8	1	8	8	0	0	0
Jaso, J, C		3	2	1	0	1	0	0	.270	(	Overba	y, 1B				5	1	1	3	0	0	4
Bartlett, SS		3	3	2	0	1	0	0	240	E	ncam	acion, 3	В			5	2	3	1	0	0	0
Brignac, R, 2B		3	1	1	1	0	1	1	264	1	Arencia	ia,C				5	3	4	9	0	0	1
Totals		33	11	9	10	7	5	11	.252	1	Totals					43	17	20	17	6	4	26

Table 2. MLB Data of Ramirez j vs left (mlb.mlb.com/stats)





NSW WARATAHS RUGBY TEAM

Figure 2. IBM's IT Platform for Sports Ecosystem (IBM Press, 2017)

As such, Big Data Analysis can be used in various ways to make timely and adequate decisions in professional sports games. IBM developed IT platform to help sports enterprises benefit from a sports ecosystem where sports owners drive engagement, performance and optimization as shown in figure 2. Australia's New South Wales Waratahs rugby team significantly reduced players' injuries by applying IBM's analytical approach, where their intensity, collisions and

fatigue levels are thoroughly examined through GPS trackers physically attached to them (IBM, 2017).

#### C. Blockchain Analysis of a Professional Sports Player

It is not yet known the use of Blockchain technology to analyze the performance of professional sports players and teams. However, as discussed above, the increasing complexity of data such as the amount of data, the psychological characteristics of individual players, social issues in Social Networks, and the emergence of users of various sports information calls for a new framework of analysis.

As Blockchain's distributed ledger technology using Big Data develops, the application of the technology can improve the performance of professional sports players and teams. Supervisors, coaches as well as professional players are benefited from the access to real-time Big Data related to the players' and teams' performance. Blockchain's distributed ledger technology will be able to provide the benefits to the interest groups, who own the distributed ledgers holding Big Data.

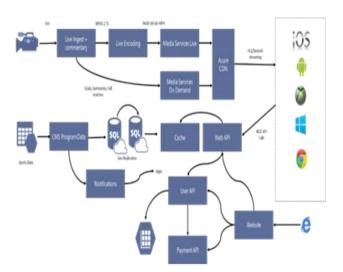


Figure 3. Microsoft Azure Platform in Sports (Azure, 2016)

The Microsoft Azure Platform shown in Figure 3 below, the blockchain structure developed by AMIS and the front channel developed by ITRI, a high profile and secure sports activity sharing platform is established.

Bravelog network, which registers "running" competitions and the profile of athletes, using Microsoft Azure, and establishing a Blockchain structure to share data can benefit professional sports industry in many ways; (1) performing complete backup and simultaneous data maintenance, (2) providing distributed replica, shared ledger and so on, which allow a low-risk and simplified process, (3) providing a player based rather than an event based enquiry structure, (4) being accessed across multiple users and multiple devices, (5) enabling players to understand their capabilities, manage their personal training journey effectively, design training plan accordingly, so as to achieve the goal of facilitating sports performance and expand the spectrum of the sports industry (PR Newswire, 2016).

# IV. DESIGN BIG DATA AND BLOCKCHAIN SYSTEM FOR PROFESSIONAL SPORTS TEAM

### A. Big Data System Design of a Professional Sports Player

The first step is to accumulate the Big Data for each player to improve the performance of each player and increase the overall winning odds of the team. First, in-game data within the season can be collected through a wide variety of sources available. Then, it is possible to select which data is most effective to increase the odds; for example, by analyzing each player's odds by an opponent team or by an individual opponent player. Accumulating and analyzing such in-game data in the time series from the past as far as possible is also suitable for analyzing the tendency of the player according to the player's condition tendency or the relative odds by opponent teams or players.

It is important to use wide range of big data, which is not only in-game data but also outside-of-game data. Thus, in this system design, information for each athlete across all areas ranging from training performance, physical check, injury history, psychological data to private information from social networks. Lastly, secondary data such as results of the analysis of accumulated big data and data compiled for the use of match strategies.

#### B. Blockchain System Design of a Professional Sports Player

By using Blockchain, data accumulated in A. Big Data System Design of a Professional Sports Player above, can be arranged in blocks by each player. The data in the block is shared through distributed ledger structure in which everyone participating in the block chain network owns a distributed ledger. Performance data of sports players contained in block can be updated on a real-time basis and then shared to the ledger holders, who are the interest groups of the data. Access information and Log History accessing the shared information is also recorded real-time.

## C. Design of Information Sharing and Spreading through Blockchaining and System Security

In this system design, Big Data and the blocks, as designed in A. Big Data System Design of a Professional Sports Player and B. Blockchain System Design of a Professional Sports Player above, are linked in chains or in times-series so all system users holding the blocks can access. Real-time data is updated using a communication protocol, and a user's information sharing can be confirmed using encryption and communication channels that support the hashing function.

As shown in the sample blockchain system of a professional team in Figure 4, block users are the members of a professional sports team — supervisor, general manager/coach, player, team doctor, recruiting manager, PR team, etc. They are given each node, which shares same critical information on the team, players, game strategies, etc. When each of the users trigger transactions, in-puts and outputs of the information linked in the chain, the transactions are approved by approvers, in this case, managers of the professional team members. The approved transactions are updated so each node still share the same updated information.

This system also enhances the information security level through access control by users as shown in the table III below. For example, a supervisor and a general manager of a team can have access to all data with no limitations. However, not all users of the data in the system can have unlimited access. Team doctors, for example, may not need to know the team's critical match strategies and each athlete's private

information, which can be open to them only when that information is necessary for medical purposes. Access of the players in the team should be limited to salaries and trade information. Recruiting managers as well as PR team should have access to only relevant information.

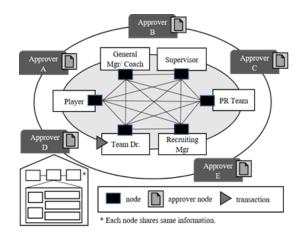


Figure 4 Sample Blockchain System Design of Professional Sports Team

Table 3. examples of access control by interest groups using the blockchain system

	Super -visor	General Mgr/ Coach	Team Dr	Player	Recruit -ing Mgr	PR Team
ACCESS CONTROL	All	All	Excl. Match Strate- gy, Private Inform- ation, etc. (unless nece- ssary)	Excl. Salary inform ation, Trade inform- ation, etc.	Match Strateg y, etc. (unless nece- ssary)	Only select- ed inform ation is avail- able for Public Rela- tions

#### D. Blockhain System Design using Hyperledger Fabric Technology

In this system design, Hyperledger Fabric blockchain technology is applied. Hyperledger Fabric technology has advantages such as strict management of accounts and selective consensus algorithm, which makes it superior to other blockchain technology for this report's purpose.

#### 1. Hyperledger Project

Hyperledger Project uses fabric code based on Apache License Version 2, as shown in the figure 5 below. This source is open software available in http://www.hyperledger.org. Originally, IBM designed this blockchain code for the purpose of their Hyperledger Project.

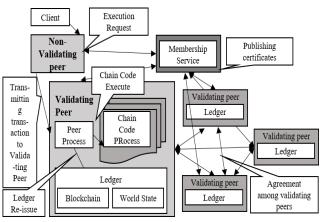


Figure 5. Transaction stream in Hyperledger
Fabric

#### 2. Permissioned Network

Hyperledger Fabric as shown in the Figure 6, is Permissioned Network. Permissioned Network is permission-based network, where not everyone has access to the network unlike Bitcoin Core or Ethereum. Under the Permissioned Network, those who intends to participate in the network or use the network needs to be registered prior to their access.

Information of the registered participants is saved and managed via the network's membership service. Membership service manages user accounts via ID registration, issuance of Enrolment Certificate (ECert), Issuance of Transaction Certificate and issuance of TLS Certificate.

#### 3. Consensus Algorithm

Hyperledger Fabric uses consensus algorithm to reach agreement among the participating nodes in the network. Under the permissioned network, participating parties reach agreement on the result of transactions through voting system. Under the Hyperledger Fabric, there are two consensus algorithm – No-Op (No operation) and Practical Byzantine Fault Tolerance (PBFT). PBFT can be properly

operated even when the computers are malfunctioning or there are errors in the network. This algorithm ensures reliability even when part of the computers in the network is lying while many computers participate in the network.

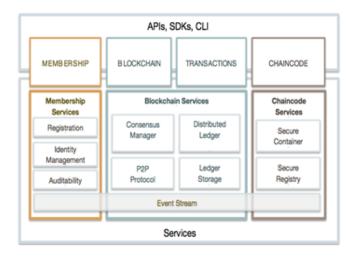


Figure 6. Transaction stream in Hyperledger Fabric (Hyperledger)

#### 4. Ledger

Data stored in the Fabric is called Ledger. The World State is a database called Key Value Store, which stores executed transaction results. When a transaction is closed, the final state is stored in the World State. For example, if A transfers money to B, the transaction result of increased A's balance and decreased B's balance is recorded.

#### 5. Chain Code

Fabric is a 'chain code' program, which is used to execute transactions in the network. Chain code consists of three stages

– Init, Invoke, and Query. Interface follows the chain code according to the package in; https://godoc.org/github.com/hyper

ledger/fabric/core/chaincode/shim. Init and Invoke updates the World State. The result is recorded in a block. Query-related block does not update new information, thus is not separately created. Fabric issues ID and distributes when the chain code is created to distinguish various chain codes in the network.

#### 6. Validation Peer

Validating Peer node makes blocks in the blockchain. Validating Peer creates transactions and validates them. Based on the transactions created by a Validating Peer is linked to a chain. The result is agreed by another Validating Peer. Agreed results is updated in the blockchain and the ledger in the World State.

#### 7. Running Hyperledger Fabric

Hyperledger Fabric's execution environment runs automatic environment using Vagrant.

#### <Source Code Clone>

```
cd %GOPATH%
mkdir src github.com hyperledger
cd src github.com hyperledger
git clone https://github.com/hyperledger-archives/fabric.
git
```

#### <Result>

```
Cloning into 'fabric'...
remote: Counting objects: 26924, done.
remote: Compressing objects: 100% (4/4), done.
remote: Total 26924 (delta o), reused o (delta o), pack-
reused 26920
Receiving objects: 100% (26924/26924), 45.26 MiB | 7.95
MiB/s, done.
Receiving deltas: 100% (18016/18016), done.
Checking out files: 100% (1164/1164), done.
```

#### 8. Writing HTML files

Request transaction can be executed after use login is done and chaincode is distributed accordingly.

When user name and password information is inserted and logged in, press 'deploy' button to deploy the chaincode into the blockchain network.

When the deploy is completed, press 'refresh' button to show the counter item list. To raise the count, choose the counter item from the list and press 'countUp' button. Create 'index.html' using the following text in the text editor. Save the HTML file in the form of UTF-8. <Counter Application Program>

```
[index.html]
<IDOCTYPE html>
<html lang="KO">
<head>
         <meta charset="UTF-8">
         <title>COUNTER</title>
         <script type="text/javascript"
         src="http://ajax.googleapis.com/ajax/libs/jquer
         y/2.1.0/jquery.min.js"></script>
         <script>
                  var url;
                  var url_host = "https://localhost:5000
                  var ccId;
                  var counterList;
                  var user_name;
                  // Log In
                  function login() {
                           url = url_host + "registrar";
                           user_name = $("#userName"
                           ).val();
                           var password - $("#password"
                           ).val();
                           var JSONdata = {
                                    "enrollId": user_nam
                                    "enrollSecret": passw
                           };
                           executeJsonRpc(url, JSONdat
                           a,
                                    funcion success(data
                                    // Log In Success
                                             console.log(
                                             "login succe
                                             ss!";
                                    function error(data)
                                    // Log In Fail
                                             console.log(
                                             "login error!
                           );
                  //Distribute
                  function deploy() {
                           url = url_host + "chaincode";
                           var JSONdata =
                           createJSONdataForCounterA
                           pp("deploy", "init", [], 1);
                           executeJsonRpc(url, JSONdat
                                    function success(dat
                                    a) {
                                             //Distributi
                                             on Success
                                             ccId = data.
                                             result.mess
                                             age;
                                             console.log(
                                             "deploy suc
                                             cess!");
                                   },
```

```
(continued from previous page)

) {

//Distrib
ution Fail
console.l
og("deplo
y error");
}

);
}
```

#### 9. Monitoring Tool

To create monitoring tool to monitor the blockchain status, create 'monitor.html' file and put the file in the directory such as 'index.html.'

Input the following text in the text editor using the UTF-8 form.

```
[monitor.html]
<!DOCTYPE html>
<html lang="ko">
<head>
         <meta charset="UTF-8">
         <title> Monitoring Tool </title>
         <script type="text/javascript" src="http://ajax.googleapis.com/ajax/libs/jque
         ry/2.1.0/jquery.min.js"></script>
         <script>
         var url_ccInfo = http://localhost:5000/chain";
         var url_prefix ="http://localhost:5000/chain/b
         locks/";
var blockNum = 1; // Block Number
         var chainHeight; // Block Height
         var stop = false;
         // Monitoring Initiation
         function startMonitor() {
                  stop = false;
                  var table = document.getElementById
                  ('list');
                  //Obtain Chain Information
                  getChainInfo(function success(data) {
                            chainHeight - data.height;
                            for (var i = 1; i <= chainHeigh
                            t; i++) {
                                     //Obtain Block Info
                                     mation
                                     getBlockinfo(data, t
                                     able);
                  }, function error (data) {};
                  //Obtain Next Block Solution Informa
                  tion Per Every Second
                  SetInterval(function(data, table) {
                  watchBlock(data, table);
                  }, 1000);
```

```
(continued from page left)
        //Monitor Stop
        function stopWatch() {
                stop = true;
                console.log("stop!");
}
        //Block Information Editing
        function insertBockRow(data, table) {
                var row = table.insertRow();
                var td = row.insertCell(o);
                td.innerHTML = blockNum++;
                var td = row.insertCell(1);
                var date = new Data(
                parseInt(data.transactions(o).timesta
                mp.seconds)*1000);
                td.innerHTML = date.toString();
                var td = row.insertCell(2);
                td.innerHTML = data.previousBlockH
                ash:
                var td = row.insertCell(3);
                td.innerHTML = data.transactions (o).
                nounce;
                var td = row.insertCell(4);
                td.innerHTML = <"input type = 'text' v
               alue = "" + JSON.stringify(data.transac
tions) + ""/>";
</head>
<body>
        <br />
        <input type="button" value="start" onclick="st
        artMonitor();"/>
        <input type="button" value="stop" onclick="sto
        pWatch();"/>
        Block<br />Number
                        TimeStamp
                        BlockHash
                        Nonce
                        Transaction
                </body>
</html>
```

When all the files created are deployed in the web browser, block information is shown.

Transaction information is shown in the Transaction Text territory. Each block is created for each transaction.

#### V. CONCLUSION

In this paper, the application of Big Data and Blockchain technology was designed to improve performance of professional sports players and teams. Real-life professional players were analyzed using Big Data. Blockchain System was

designed so the interest groups of professional sports teams can utilize Big Data analysis. The security level by users, thus, is determined in terms of access to data, reading, writing, and executing.

#### VI. REFEREENCES

C. Yang, X. Chen, Y. Xiang, Feb 2018, 'Blockchain-based publicly verifiable data deletion scheme for cloud storage,' J. Netw. Comput. Appl., vol. 103, pp. 185–193.

https://hyperledger-fabric.readthedocs.io

IBM. IBM predictive analytics reduces player injury and optimises team performance for NSW Waratahs rugby team.

http://www03.ibm.com/press/au/en/pressrelease/42613.wss

- J. Gantz, D. Reinsel, 2011, 'Extracting value from chaos,' *IDC iView*, vol. 1142, pp. 1–12.
- J. H. Park, H. J. Hyuk, Aug 2017, 'Blockchain Security in Cloud Computing: Use Cases, Challenges, and Solutions,' *SYMMETRY-BASEL*, vol. 9, no. 8, pp. 1–13.
- J. Li, S. Y. Shin, H. C. Lee, Sep 2017, 'Construction of Scientific Impact Evaluation Model Based on Altmetrics,' J. Inf. Commun. Converg. Eng., vol. 15, no. 3, pp. 165-169.
- L. Zhe, K.-K. Raymond, M. Zhao, Aug 2017, 'Practical-oriented protocols for privacy-preserving outsourced big data analysis: Challenges and future research directions,' *Comput. Secur.*, vol. 69, pp. 97–113.
- Microsoft Azure. Next Generation Sports Network launches global sports streaming, powered by Azure Media Services. https://azure.microsoft.com/en-gb/blog/next-generation-sports-network-launches-global-sports-streaming-powered-by-azure-media-services/
- PR Newswire. 04 January 2016, A New Milestone of Microsoft Azure: The First Sports Blockchain "BraveLog" https://www.prnewswire.com/ news-releases/a-new-milestone-of-microsoft-azure-the-first-sports-blockchain-bravelog-available-online-300385433.html
- Y. Hwang, K. Kim, O. Kwon, I. Moon, G. Shin, J. Ham, J. Park, Dec 2017, 'Analyzing Box-Office Hit Factors Using Big Data: Focusing on Korean Films for the Last 5 Years,' *J. Inf. Commun. Converg. Eng.*, vol. 15, no. 4, pp. 217–226.