### Distributed Blockchain-Based Data Protection Framework for Modern Power Systems Against Cyber Attacks

Author – Gaoqi Liang, Steven R. Weller, Fengji Luo, Junhua Zhao, Zhao Yang Dong Published in IEEE Transactions on Smart Grid. VOL.10, No. 3, MAY 2019

present

2019.05.14

**Minjeong Cho** 



#### IEEE Transactions on Smart Grid

ISSN: 1949-3053 eISSN: 1949-3061 IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

Go to Journal Table of Contents Printable Version

TITLES ISO: IEEE Trans. Smart Grid JCR Abbrev: IEEE T SMART GRID

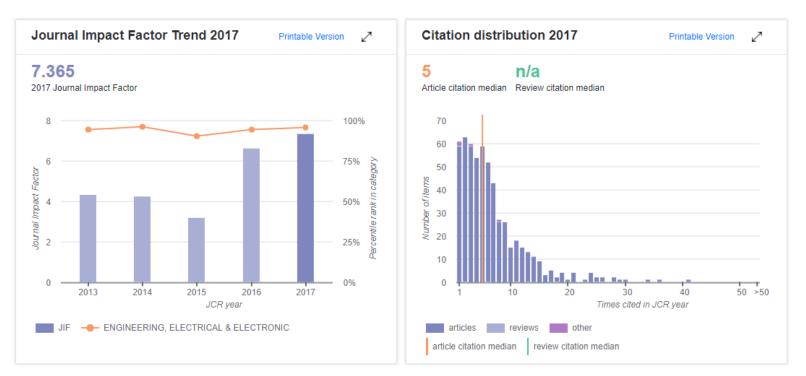
CATEGORIE S

ENGINEERING, ELECTRICAL & ELECTRONIC - SCIE LANGUAGES English

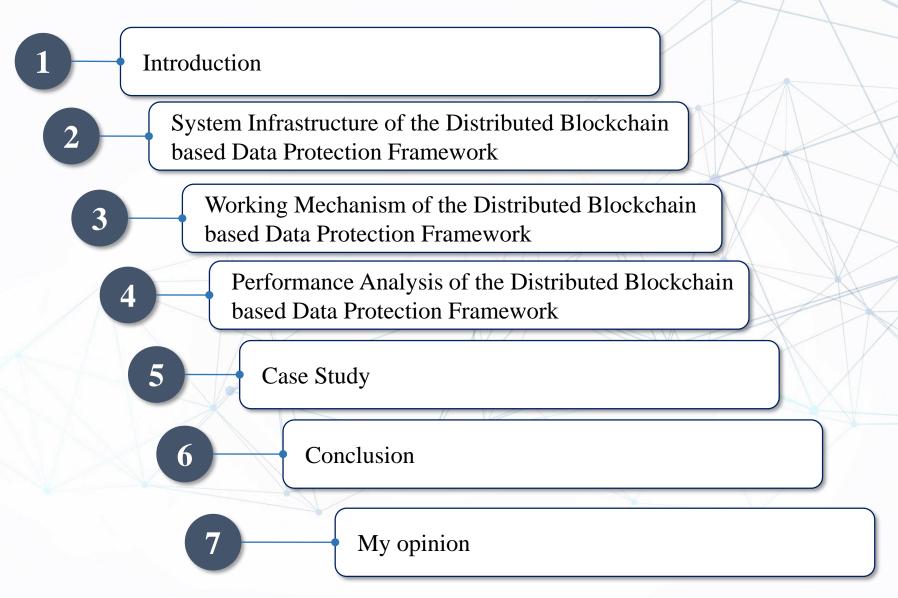
PUBLICATION FREQUENCY 6 issues/year

Current Year All years

The data in the two graphs below and in the Journal Impact Factor calculation panels represent citation activity in 2017 to items published in the journal in the prior two years. They detail the components of the Journal Impact Factor. Use the "All Years" tab to access key metrics and additional data for the current year and all prior years for this journal.



### Contents



### 1. Introduction

Modern power systems have experienced a profound evolution to facilitate social development

While this technology trend on the one hand provides new opportunities to optimize the energy efficiency of grid, it also imposes significant requirements and challenges on the **robustness, efficiency, and security** of the underlying information infrastructure

Due to the deep integration of both cyber and physical resources, attacks from the cyber layer have the potential to mislead decision-making the control center and cause system disturbances, financial loss, or even more serious consequences.

As a representative cyber-attack, the false data injection attack(FDIA) manipulates system data to mislead the control center without being detected by the bad data detection module

In this sense, data vulnerability has become an unneglectable issue, as evidenced by malicious events caused by cyber-attacks, a recent high-profile example of which was the 2015 Ukraine blackout

### 1.Introduction

Main contribution

- 1. The proposed framework substantially increases the self-defensive capabilities of modern power systems against data manipulation by cyber attackers.
  - $\checkmark$  In conventional power systems, an attack is deemed successful if cyber attackers tamper with
    - meter measurement data locally
    - replace data packages transmitted to the control center via a communication channel
    - hacks into control center.
  - ✓ In the proposed framework, an attack does not result in a successful manipulation unless an attacker tampers with (or replaces)
    - data packages on a majority of channels
    - hacks into sufficient meters to manipulate data.
- 2. The proposed framework is consensus-based, and exploits particular characteristics of the power grid environment

#### 2. System Infrastructure of the Distributed Blockchain based Data Protection Framework

Typically three basic processes for the SCADA module in modern power system

- Data gathering at remote terminal units
- Plaintext transmission via a communication channel to the control center
- Information storage in the control centre

Current information-gathering and storage mechanism provides centralize management

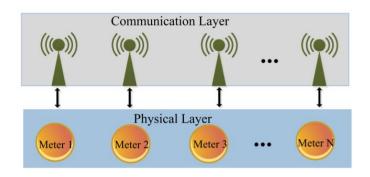
high risks of data being manipulated by cyber attack

Proposed framework provides a distributed information gathering and storage mechanism



### 2. System Infrastructure of the Distributed Blockchain based Data Protection Framework A. Reconfigured SCADA Network

Some system infrastructure must be updated or replaced to facilitate the working mechanism of the proposed framework

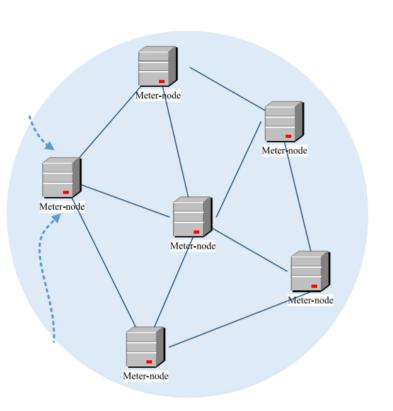


The overall power system layers are as usual

- Data still collect real-time measurements from the grid including voltage, current, real and reactive power flow, breaker status, transformer tap position, and so forth
- The communication layer is isolated from the Internet
  - meters/sensors distributed geographically

But meter is assembled by ①data collection device, ②signal sender, ③signal receiver and ④data process device

2. System Infrastructure of the Distributed Blockchain based Data Protection Framework A. Reconfigured SCADA Network



Meter/sensors acts as node

The graph corresponding to the meter-node network is connected

there is communication path linking each distinct pair of nodes

Meter-node network can be considered as a private blockchain network

Only meters/sensors which are authorized by the grid can perform data acquisition function

Interactions among the nodes in the network are automatically performed based on a certain consensus mechanism(Without human interaction)

### 2. System Infrastructure of the Distributed Blockchain based Data Protection Framework B. Key Features of Meters

In order to interact with each other through the proposed blockchain framework, each meter needs to be possess functional features which are not common in today's widely deployed meters

**Required features** 

- Each meter is identified by a unique address
- Each meter is equipped with specific software to support the generation of a public key and private key
- Each meter is equipped with RAM, computational hardware, data collection device, signal sender, signal receiver and data process device
- Meter are capable of communicating with each other though wired or wireless communication channels

#### 3. Working mechanism of the distributed blockchain data protection framework

In the proposed framework, all collected data are eventually stored in a ledger in a form of connected blocks which exists in distributed form in each meter's memory

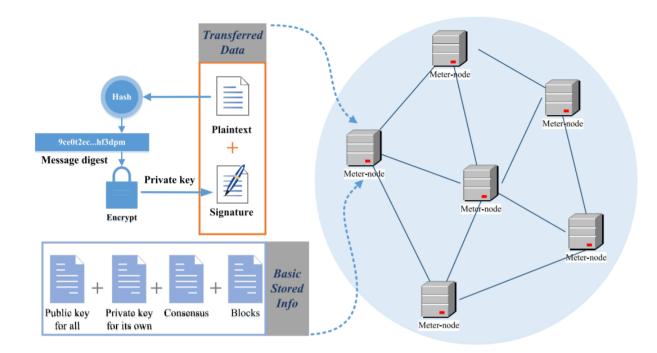
Before storage, some procedures are necessary to guarantee data accuracy

- Data broadcast
- Data verification via voting mechanism
- Data content accumulation in block
- Mining process
- Verification the mining result via voting mechanism
- Distributed ledger synchronization

Main working mechanism

- Data transmission
- Verification
- Storage

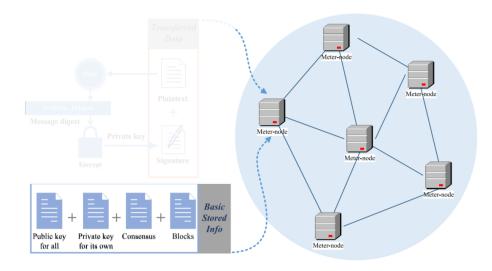
### 3.Working mechanism of the distributed blockchain data protection framework A. Data Encryption and Broadcast



#### **Data Encryption and Broadcast Process**

Data within each meter-node is comprised of basic stored information and transferred data

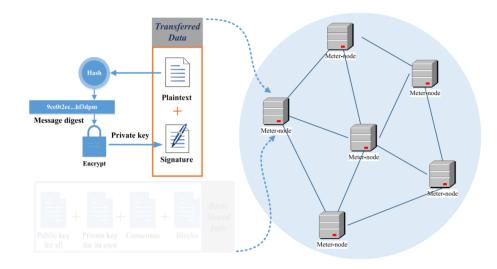
### 3. Working mechanism of the distributed blockchain data protection framework A. Data Encryption and Broadcast



#### Basic Stored Info

- 1. Public key for all meter-nodes
  - Public key is node's main accessible information that is publicly available in the meter-node network.
- 2. Private key for its own
  - node's private information that is used to validate a node's identity and operations that it may perform
- 3. Preset consensus
- 4. Accumulated blocks

### 3. Working mechanism of the distributed blockchain data protection framework A. Data Encryption and Broadcast

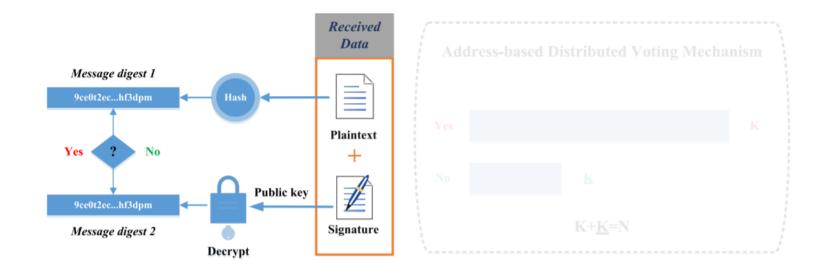


#### Transferred Data

- 1. Plaintext
- 2. Signature
  - Newly collected plaintext is processed using a secure hash algorithm(SHA), generating a message digest.
  - The private key of each node is used to encrypt the message digest of that node
  - Forming a digital signature which can be decrypted using its public key

The transferred data is then broadcast to all other meter-nodes via the communication network

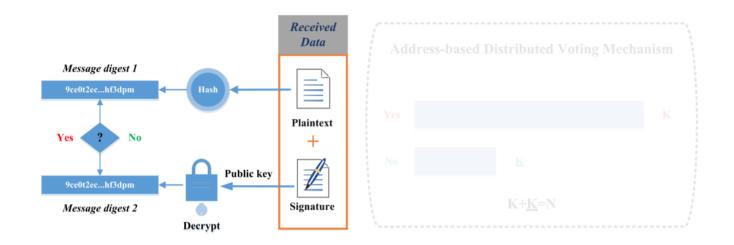
## 3. Working mechanism of the distributed blockchain data protection framework B. Data Decryption and Verification



#### **Data Decryption and Verification Process**

All meter-nodes which receive broadcast information need to decrypt the received data and verify the results

## 3. Working mechanism of the distributed blockchain data protection framework B. Data Decryption and Verification



- 1. The receiver hashes the received plaintext *into message digest 1*
- 2. Decrypt message digest 2 from the digital signature by using the sender's public key
  - ① If message digest 1 equals message digest 2, the received information is successfully verified
  - ② Otherwise the received data is considered as false

data integrity and consistency issues exist in the broadcasting process

All nodes use an address-based distributed voting mechanism

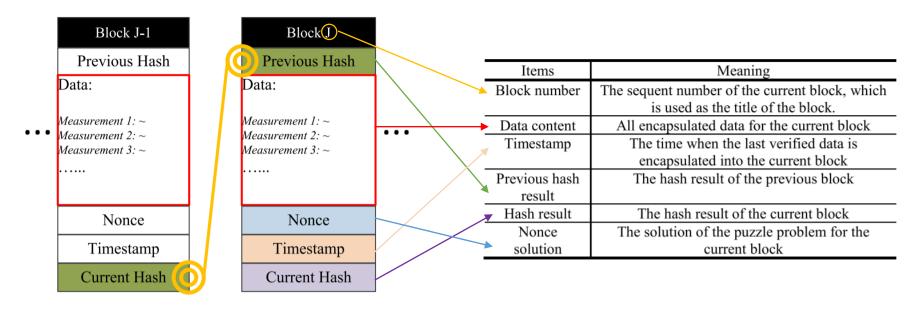
- Each node has precisely one chance to verify the integrity and consistency of the received data
- → Only once positive agreement is reached among nodes is the data recognized as correct

## 3. Working mechanism of the distributed blockchain data protection framework B. Data Decryption and Verification



#### Criterian for data accept

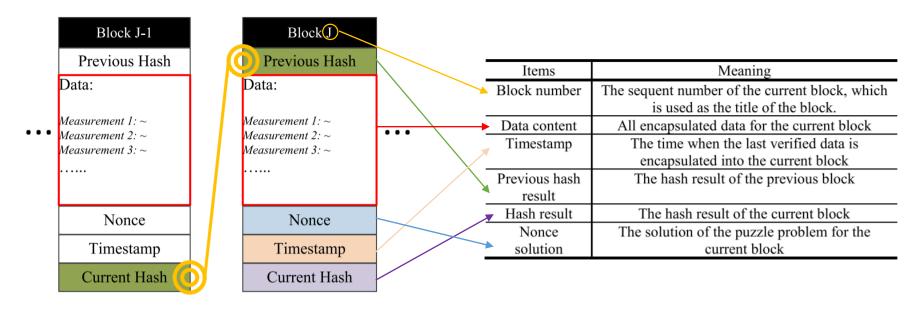
- $\implies \frac{K}{N} > \tau \text{ (N : meter-node network,}$ 
  - K : most voted
  - $\underline{\mathbf{K}}$ : other
  - $\tau$  : threshold whose value must be strictly greater than 50%)



#### **Block Contents and Chain Connections.**

- Block number
- Data content
- Timestamp
- Previous hash result
- Hash result
- Nonce solution

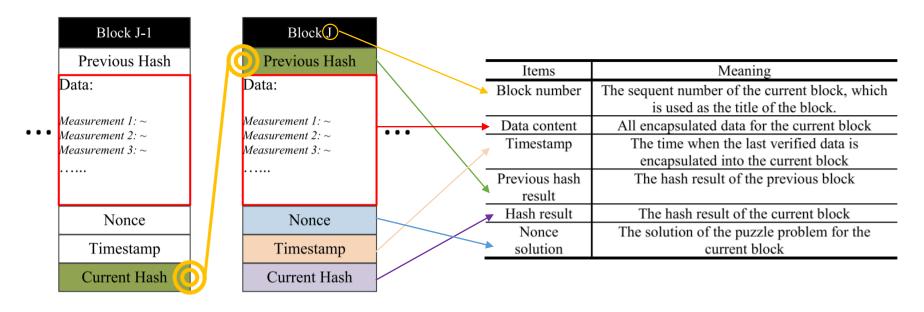
#### Meaning of the Attritbutes



#### **Block Contents and Chain Connections.**

- 1. Pre-processing : S = b + d + t + hp + nonce
  - 1. b:blocknumber
  - 2. d : data content
  - 3. t: time point
  - 4. hp : previous hash result
  - 5. Nonce : random number

#### Meaning of the Attritbues



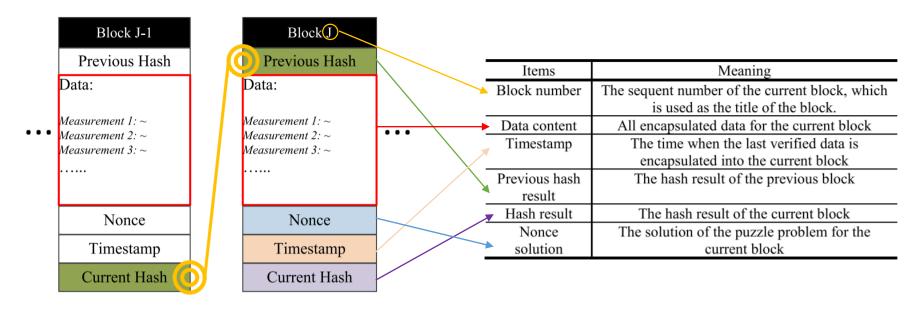
#### **Block Contents and Chain Connections.**

#### Meaning of the Attritbues

2. FinalHash = hash(SHA256, hash(SHA256, S))

The puzzle problem is to find the appropriate nonce value to make the FianlHash value less than a given target T *FinalHash*  $\leq T$ 

- 3. Some nodes can operate as miners by attempting to sole the puzzle problem independently
- 4. Once the first miner finds the nonce, it broadcasts the value to other nodes to let them check whether the solution is correct by validating where it satisfies  $FinalHash \le T$
- 5. Address-based distributed voting mechanism is used again to vote on the verification result



#### **Block Contents and Chain Connections.**

#### Meaning of the Attritbues

- Mining is a competition among all miners
- There is no reward as an incentive for miner who solves the puzzle problem first
- All nodes are strictly driven by the consensus
- All miner behaviours are pre-programed and automatically generated

3. Working mechanism of the distributed blockchain data protection framework D. Consensus Mechanism

#### Setting of Public/Private Key Update Frequency

**Block Generation Frequency** 

**Miner Selection Method** 

**Release of Meter's Memory Periodically** 

#### 3. Working mechanism of the distributed blockchain data protection framework D. Consensus Mechanism

#### **1.Setting of Public/Private Key Update Frequency**

- If Public key and private keys are stolen by an adversary, it would be challenging for the network for the network to prevent data from being manipulated by cyber attackers
- Regular update/replacement on key information is therefore an effective method of enhancing security
- **\*** Key update Frequency should satisfied  $\Psi < t_K$
- ◆ t<sub>i</sub> is is estimated average time for the attacker to steal the public and private key of i-th node(solted in increasing order) t

$$_{i}, i = 2, 3, ..., N$$
 where  $t_{i} > 0,$ 

$$t_{i-1} < t_i$$

- Minimum number of stolen paris of public and private keys on nodes in order to tamper  $K > \tau \cdot N, K = ceil(\tau \cdot N)$ , (ceil denotes round up)
- Required time for an attacker to steal key information form all K nodes

(scenario in which an attacker only has the capability of stealing a single key at time)

$$t_K \leq \bar{\Psi} \leq \sum_{i=1}^K t_i$$
  $t_K = \max\{t_i\}, i = 1, 2, \dots, K.$ 

(scenario in which an attacker only has the capability of stealing key information from K nodes simultaneously)  $t_K \leq \bar{\Psi}$ 

### 3. Working mechanism of the distributed blockchain data protection framework D. Consensus Mechanism

#### 2. Block Generation

If one block accumulates excessive measurement data, the process could take sufficiently long by adversely impacted Too frequency mining is computational burden for the blockchain system

Strategy 1. Generating Block by Fixed Time

 $\alpha < \Phi$   $\beta \cdot floor\left(\frac{\Phi}{\alpha}\right) \ge N$ Number of meter-node *N*, time interval of block generation  $\alpha$ , averge number of measured data itmes in each block  $\beta$ , system state estimation time interval  $\Phi$ , rounding down function *floor* 

Strategy 2 . Generating Blocks by Fixed Size

$$\bar{\alpha} < \Phi$$

$$\bar{\beta} \cdot floor\left(\frac{\Phi}{\bar{\alpha}}\right) \ge N$$
Number of meter-node *N*, average time interval of block generation  $\bar{\alpha}$ , block size  $\bar{\beta}$ , system state estimation time interval  $\Phi$ , rounding down function *floor*

#### 3. Working mechanism of the distributed blockchain data protection framework D. Consensus Mechanism

#### 3. Miner Selection

Miner must be equipped with substantial computational capability But it requirement potentially implied high investigation costs

#### Strategy 1. Pre-Specified Nodes As Miners

Some nodes are pre-specified to act as miners, and are responsible for solving the puzzle problem

- Pros : compromising between the mining efficiency and computational hardware investment is possible
- Cons : pre-specified miners could become the targets of cyber-attacks

#### Strategy 2. Randomly Selected Nodes As Miner

The computational hardware configurations of all the nodes are same, but not all nodes are required to act as miners

- Pros : more secure
- Cons : computation hardware configuration of all the nodes are same greater investigation in hardware
  - : complex as each time the miners need to be re-selected

#### 4. Release of Meter's Memory Periodically

- With continuous operation of the system, the blockchain ledger will become progressively larger
- The data content of the blocks needs to be backed up and meter memory released periodically ٠

### 4. Performance Analysis of the Distributed Blockchain base Data Protection Framework A. Blockchain Technology Innovation and Comparison

Items	Blockchain in Bitcoin System	Blockchain in the Proposed Framework
Network	Public	Private
Transaction initiator	Human intervention	Completely automatic
Transaction content	Money	Collected measurement
Transaction relationship	Continuously, related	Independent, unrelated
Checking historical blocks prior to the voting process	Required	Unnecessary
Chain connection speed	Approximately 7 transactions per second [37]	Much faster
Reward to node	Yes	No
Double-spending attack	A threat	Not exist
51% attack	Difficult	Difficult but threshold adjustable

**Technology Comparison** 

4. Performance Analysis of the Distributed Blockchain base Data Protection Framework B. Potential Disadvantages and Practical Challenges

Items	Disadvantages		Challenges
Timeliness	Upgrade Sensing devices		Cost
	/replacement	Communication	vs. benefit
		networks	
Security	Majority	Sensors	Technology
	Manipulated	Communication channels	development
Redundancy	Information	Distributed data storage	Defending
	disclosure		strategy

#### **Potential Disadvantages and Practical Challenges**

• Timeliness

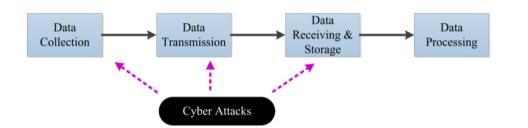
To balance the necessary investment in upgrades or replacement with the benefit from enhanced security.

• Security

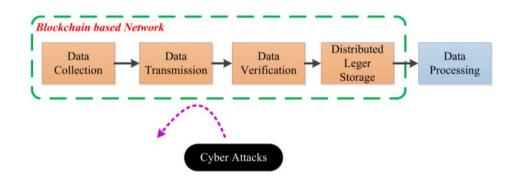
The proposed framework is based on the mechanism of the majority rule, i.e., geographically distributed sensors—hence many communication channels—greatly increases the difficulty faced by cyber attackers in manipulating sensors/channels so as to reach a false agreement.

• Redundancy

since each one of the registered sensors/meters in the network has a record of all nodes' measured data during some period of time. An attacker may therefore read all distributed stored data by hacking into a single sensor/meter.



General procedure for existing data communication



General procedure for blockchain-based data communication

• General procedure for existing data communication

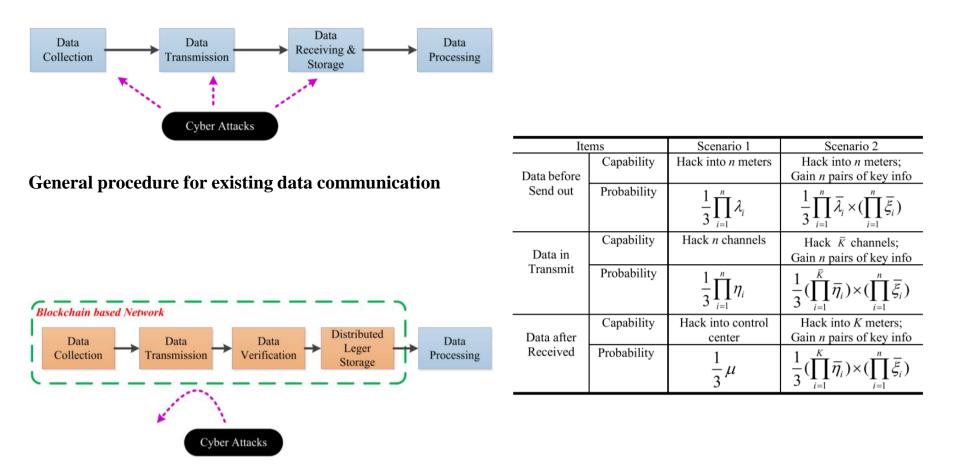
Cyber-attackers may manipulate data after it is collected, during data transmission, or when data is received and stored in control center.

Three forms of manipulations are independent

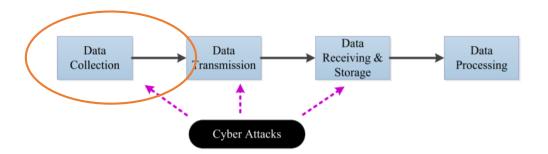
• General procedure for blockchain-based data communication

Cyber-attackers may manipulate data after it is collected (but prior to broadcast), or when data is transmitted to all other nodes via communication channels, or after data has been received at nodes (but prior to the data verification stage) in such a way to reach a false agreement through the voting mechanism

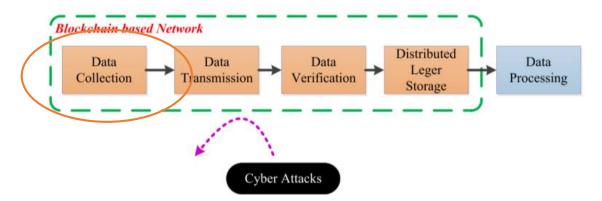
Probability for attackers to steal each meter's key information is independent



General procedure for blokchain-based data communication



General procedure for existing data communication



General procedure for blokchain-based data communication

Ite	ms	Scenario 1	Scenario 2
	Capability	Hack into n meters	Hack into n meters;
Data before			Gain <i>n</i> pairs of key info
Send out	Probability	$\frac{1}{3}\prod_{i=1}^n \lambda_i$	$\frac{1}{3}\prod_{i=1}^{n}\overline{\lambda}_{i}\times(\prod_{i=1}^{n}\overline{\xi}_{i})$

Scenario 1

n: first n meters that attacker needs to manipulate

 $\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots, \lambda_N$  : probability for attackers to hack into each meter in independent

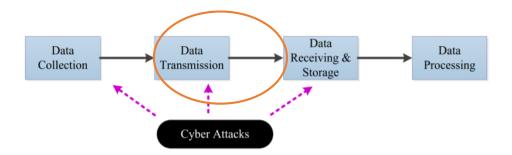
 $: 0 \le \lambda_{n} \le 1, i = 1, 2, ..., n, ..., N$ 

#### Scenario 2

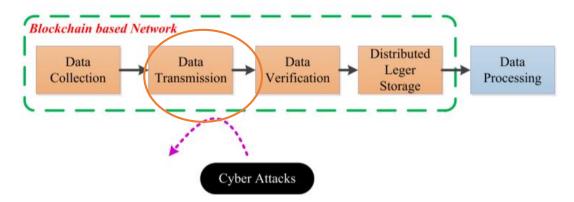
n : first n meters that attacker needs to manipulate

 $\overline{\lambda}_1, \overline{\lambda}_2, \dots, \overline{\lambda}_n, \dots, \overline{\lambda}_N$ : probability for attackers to hack into each meter is independent :  $0 \le \overline{\lambda}_n \le 1, i = 1, 2, \dots, n, \dots, N$  $\xi_1, \xi_2, \dots, \xi_n, \dots, \xi_N$ : probability for attackers to steal each meter's key information is i

 $\xi_1, \xi_2, \dots, \xi_n, \dots, \xi_N$ : probability for attackers to steal each meter's key information is independent :  $0 \le \xi_n \le 1, i = 1, 2, \dots, n, \dots, N$ 



General procedure for existing data communication



General procedure for blockchain-based data communication

Ite	ems	Scenario 1	Scenario 2
Data in	Capability	Hack <i>n</i> channels	Hack $\overline{K}$ channels;
			Gain <i>n</i> pairs of key info
Transmit	Probability	$rac{1}{3} \prod_{i=1}^n \eta_i$	$\frac{1}{3}(\prod_{i=1}^{\bar{K}}\bar{\eta}_i)\times(\prod_{i=1}^n\bar{\xi}_i)$

Scenario 1

n: first n channels that attacker needs to manipulate

 $\eta_1, \eta_2, \dots, \eta_n, \dots, \eta_N$ : probability for attackers to replace data package from the remote to control cetre for all meters

 $: 0 \leq \eta_n \leq 1, i = 1, 2, \dots, n, \dots, N$ 

#### Scenario 2

n: first n meters that attacker needs to manipulate

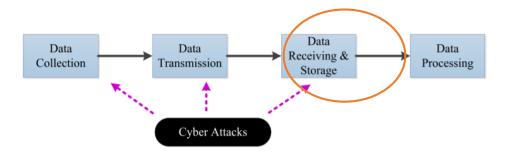
 $\overline{K}$  : Celi(N(N-1)/2 ·  $\tau$ )

 $\overline{\eta}_1, \overline{\eta}_2, \dots, \overline{\eta}_n, \dots, \overline{\eta}_{\overline{K}}$ : probability for attackers to hack into each channels is independent

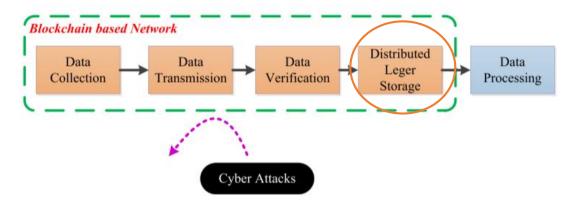
$$0 \leq \overline{\eta}_n \leq 1, i = 1, 2, \dots, n, \dots, \overline{K}$$

 $\xi_1, \xi_2, \cdots, \xi_n$ ,  $\cdots, \xi_N$ : probability for attackers to steal each meter's key information is independent

 $: 0 \leq \xi_n \leq 1, i = 1, 2, \dots, n, \dots, N$ 



General procedure for existing data communication



General procedure for blockchain-based data communication

Ite	ems	Scenario 1	Scenario 2
	Capability	Hack into control	Hack into K meters;
Data after		center	Gain <i>n</i> pairs of key info
Received	Probability	$\frac{1}{3}\mu$	$\frac{1}{3}(\prod_{i=1}^{K}\overline{\eta}_{i})\times(\prod_{i=1}^{n}\overline{\xi}_{i})$

Scenario 1

 $\boldsymbol{\mu}\,$  : probability for attackers to hack into centre

#### Scenario 2

n : first n meters that attacker needs to manipulate

 $K: celi(\tau \cdot N)$ 

 $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_n, \dots, \bar{\lambda}_N$ : probability for attackers to hack into each meter in independent

$$0 \leq \overline{\lambda}_n \leq 1, i = 1, 2, \dots, n, \dots, N$$

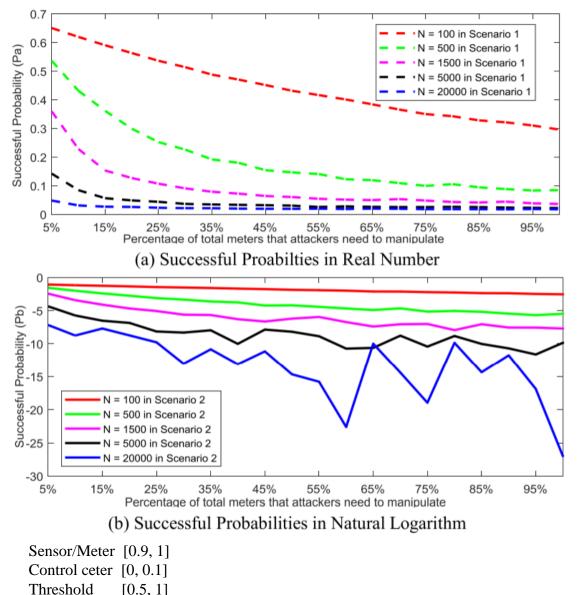
 $\xi_1, \xi_2, \dots, \xi_n, \dots, \xi_N$ : probability for attackers to steal each meter's key information is independent :  $0 \le \xi_n \le 1, i = 1, 2, \dots, n, \dots, N$ 

Scenario 1: 
$$P_a = \frac{1}{3} \left( \prod_{i=1}^n \lambda_i + \prod_{i=1}^n \eta_i + \mu \right)$$

Scenario 2

$$P_{b} = \frac{1}{3} \left[ \prod_{i=1}^{n} \overline{\lambda}_{i} \times \left( \prod_{i=1}^{n} \overline{\xi}_{i} \right) + \left( \prod_{i=1}^{\overline{K}} \overline{\eta}_{i} \right) \times \left( \prod_{i=1}^{n} \overline{\xi}_{i} \right) + \left( \prod_{i=1}^{K} \overline{\lambda}_{i} \right) \times \left( \prod_{i=1}^{n} \overline{\xi}_{i} \right) \right]$$

Items		Scenario 1	Scenario 2
Data before	Capability	Hack into <i>n</i> meters	Hack into <i>n</i> meters; Gain <i>n</i> pairs of key info
Send out	Probability	$rac{1}{3}\prod_{i=1}^n \lambda_i$	$\frac{1}{3}\prod_{i=1}^{n}\overline{\lambda_{i}}\times(\prod_{i=1}^{n}\overline{\xi_{i}})$
Data in	Capability	Hack <i>n</i> channels	Hack $\overline{K}$ channels; Gain <i>n</i> pairs of key info
Transmit	Probability	$rac{1}{3}\prod_{i=1}^n\eta_i$	$\frac{1}{3}(\prod_{i=1}^{\overline{K}}\overline{\eta}_i) \times (\prod_{i=1}^n \overline{\xi}_i)$
Data after	Capability	Hack into control center	Hack into <i>K</i> meters; Gain <i>n</i> pairs of key info
Received	Probability	$\frac{1}{3}\mu$	$\frac{1}{3}(\prod_{i=1}^{K}\overline{\eta}_{i})\times(\prod_{i=1}^{n}\overline{\xi}_{i})$



#### Monte Carlo simulation experiments

- Each variable is randomly chosen in that range for each experiment
- The simulation result shown is 1000 random trials on averages

#### N: 100,500,1500, 5000, 20000

n : increases uniformly from 5% of value N to 100 % value N with the rate at 5% for each pairs of the experiment

The largest successful attack probability for scenario 1 and scenario 2 exist in the case of manipulating 1% of corresponding N value

: (65.07%, 34.52%), (53.64%, 21%), (35.98%, 8.71%), (14.25%, 1.29%), (4.85%, 0.0078%)

#### 5.Case Study

IEEE 118-bus system

IEEE-118 benchmark system : 118 nodes, 186 branch

Each nodes deploys a meter

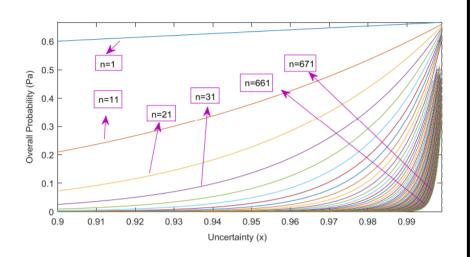
Each branch deploys a breaker

Each branch deploys two meters

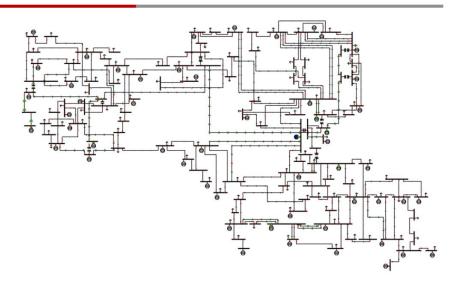
Attack : FIDA

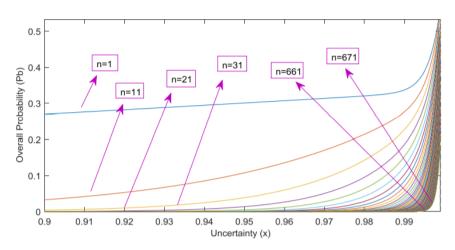
Sensor/Meter hack probability [0.9, 0.999]

Control cetre hack probability 0.001



Successful Attacking Probabilities in existing framework





Successful Attacking Probabilities in Proposed framework

### 6.Conclusion

This paper proposes a distributed blockchain-based data protection framework for enhancing the data security of modern power system against cyber-attacks

The proposed framework substantially enhances the self-defensive capabilities of power systems against cyber-attack by harnessing the distributed security features of blockchain technology first employed in the bitcoin crypto-currency

The proposed framework represents a promising new direction in cyber-security for modern power systems

The proposed framework present an evaluation of proposed framework against cyber-attack

Improvements in the underlying blockchain technology, including improvement of blocks' connection speed, acceleration of reliability and security, reduction of investment and risk, are expected to benefit blockchain-based applications

In future research, authors will consider further refinement of the consensus algorithm, and perform an assessment of associated software and hardware investment cost vs. benefits.

### 7. In my opinion

Proposed network doesn't need to using POW consensus protocol

• Proposed network is private network

Changing meters in distributed area is unrealistic.

- It spend a lot of money.
- In this paper meter should have ability to computation public key system. Public key system is known as need very high computational ability.
- many time in SCADA system required continuous operation without stopping but for changing meters, SCADA system should be stopped.

Priority of SCADA

- Even in standard document SCADA system priority is known as availability > integrity > confidentiality
- System should proof availability

Real time problem

• SCADA system is presupposing acting as real time, but in proposed framework they have too many additional element

# Thank you