Liang WANG, Shunjiu HUANG, Lina ZUO, Jun LI, Wenyuan LIU, 2023. RCDS: a right-confirmable data-sharing model based on symbol mapping coding and blockchain. *Frontiers of Information Technology & Electronic Engineering*, 24(8):1194-1213. https://doi.org/10.1631/FITEE.2200659

RCDS: a right-confirmable data-sharing model based on symbol mapping coding and blockchain

Key words: Data right confirmation; Symbol mapping coding; Blockchain; Data sharing; Traitor tracing; Access control

Corresponding author: Shunjiu Huang
E-mail: sjhuang1120@stumail.hbu.edu.cn
ORCID: https://orcid.org/0000-0003-1121-1842

Motivation

1. The purpose of data right confirmation (DRC) is to protect the legitimate ownership and usage rights of the shared data. The research on DRC is of significant importance on improving data security in many data-sharing fields.
2. Existing DRC schemes are designed for specific types of data. When dealing with different types of data content, they are not as effective as expected.
3. Although some methods of combining data watermarking and the blockchain addressed DRC problems to some extent, they are still not that practical in situations that require high data precision and critical access control.

Urgently, new right-confirmable data-sharing models need to be further studied.

Method

Symbol mapping coding (SMC) works by dividing the byte sequence of a data object into symbols and recoding them using a generated symbol mapping table (SMT).

Table 2 Schematic symbol mapping table

Symbol	Plain code	Hidden code
A(0x41)	0x0041	W(0x57)
	0xE410	O(0x4F)
E(0x45)	0xE451	M(0x4D)
	0xE452	R(0x52)
I(0x49)	0x004C	*(0x2A)
	0xE4C1	L(0x4C)
O(0x4F)	0xE4F0	D(0x44)
	0x004F	$(0x24)
U(0x55)	0x004D	!(0x21)
	0x0056	P(0x50)

RCDS integrates SMC into a blockchain system, thereby enabling credible DRC in data sharing.

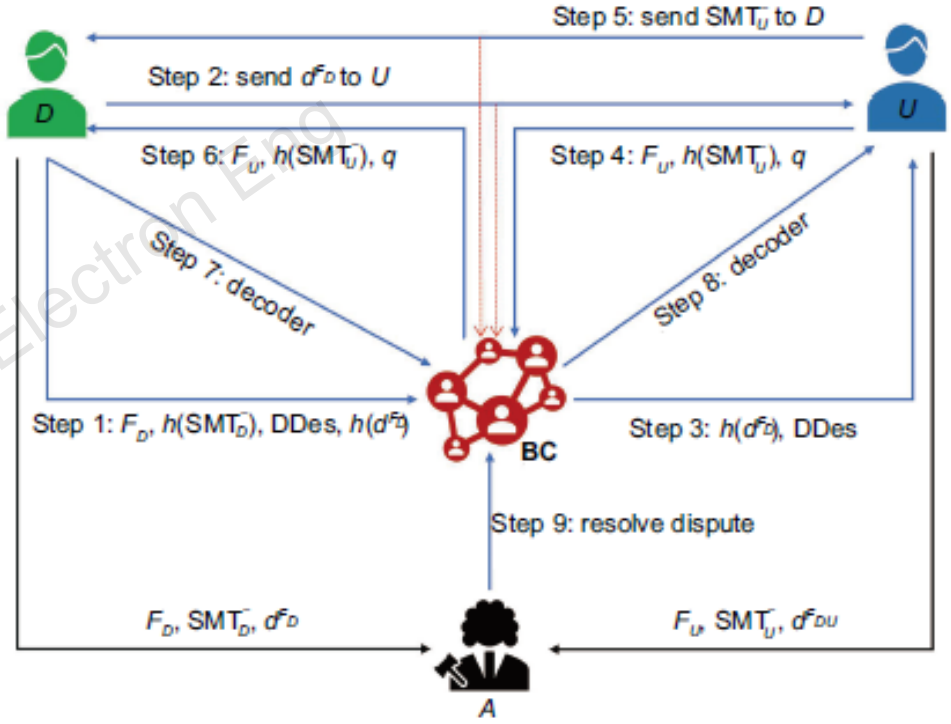


Fig. 1 Overview of RCDS

D: distributor; U: authorized user; A: arbiter; BC: blockchain; d: data object; d^F: data object with fingerprint F embedded; F_X: fingerprint of user X; SMT_X: symbol mapping table (SMT) of X; $\overline{\mathrm{SMT}_X}$: public part of SMT_X; q: data query; d^*: output of q through encapsulated decoder ⊗

Method

In RCDS, with the help of detailed description of data query DDes, a more credible access control strategy can be implemented.

Data type			Storage location	Data constraint	Data feature	Access policy
Relational data			Database	Integrity constraints	Attributes, dimensions, etc.	...
File data	Single media data	Text data	Data path	Text id	Text features	...
		Image data	Data path	Image id	Image features (color, attribute, etc.)	...
		Audio data	Data path	Audio id	Audio features (sound wave, etc.)	...
		Video data	Data path	Video structure (lens, etc.)	Video features (static features, etc.)	...
	Multimedia data		Content information (file type, entity semantion, implication relationship, etc.)	Feature information (attributes, entity similarity, etc.)	Spatial-temporal information (topological relationship, space calculation, etc.)	...

Fig. 2 DDes format

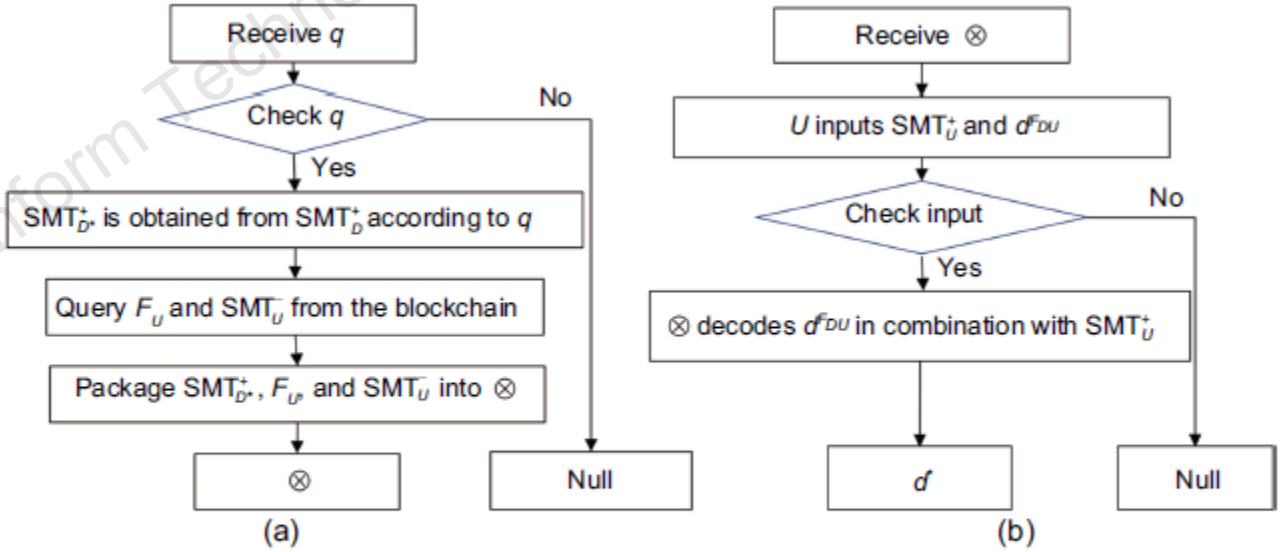


Fig. 6 Access control: (a) D phase; (b) U phase

q: query statement; d^F: data object with fingerprint F embedded; F_X: fingerprint of user X; SMT_X^-: public part of SMT_X; SMT_X^+: private part SMT_X; SMT: symbol mapping table; ⊗: encapsulated decoder

Method

The transaction chain data structure is designed to customize the blockchain data model. Doing so, the data-sharing records can be correctly collected.

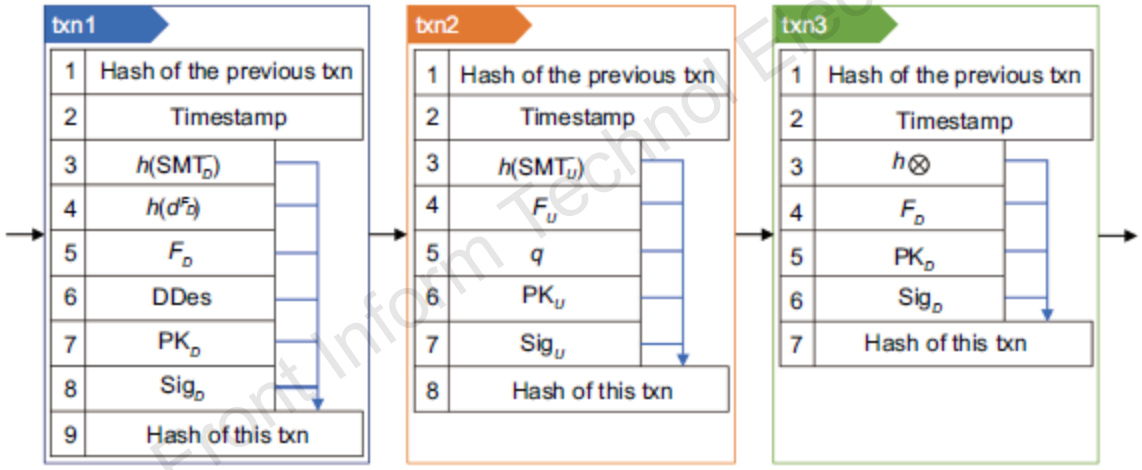


Fig. 8 Transaction chain over blockchain

txn: transaction; PK_X: public key of user X; Sig_X: signature of X; F_X: fingerprint of X; DDes: data description; d^F: data object with fingerprint F embedded; q: query statement; $h(SMT_X^-)$: digest of the public part SMT_X; $h\otimes$: digest of a decoder

Method

By optimizing parameters, the security of fingerprint identification is enhanced.

$$\begin{cases} 2|S| \le 2^{8\rho}, \\ |S| = \dfrac{|d|}{\theta}, \\ \rho|S| = |d^F|, \\ |d^F| \ge \lambda|F|, \\ \mathrm{SMT}^{-} \le \tau. \end{cases}$$

$$\left\lceil \frac{|d|}{\tau}\left(|F| + \frac{\log_2|d|+1}{8}\right)\right\rceil \le \theta \le \left\lfloor \frac{\rho|d|}{\lambda|F|}\right\rfloor.$$

$$\lceil \gamma|S| \rceil \le \varphi \le |S|.$$

$$\frac{|F|}{|S|}\lceil \gamma|S| \rceil \le \eta \le |F|.$$

The accuracy of RCDS in identifying fingerprints from data

The robustness of RCDS's fingerprint identification when facing attacks

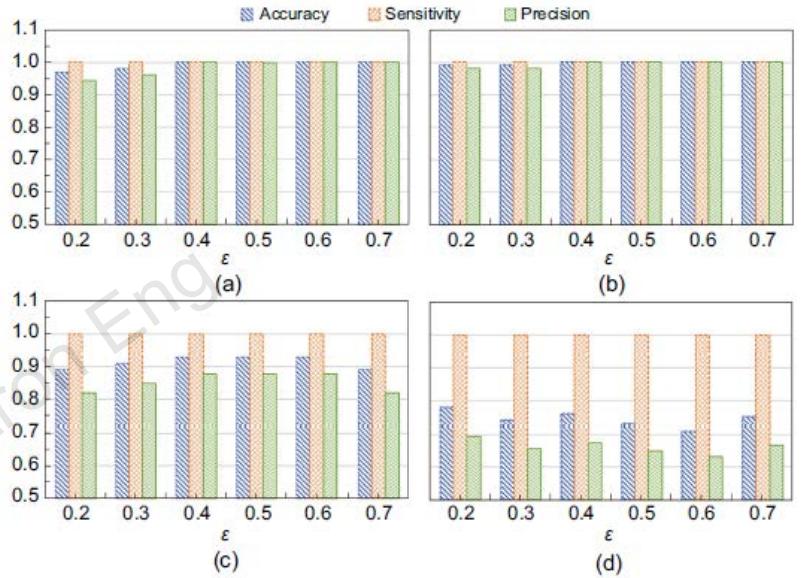


Fig. 9 Correctness of RCDS: (a) ς=10; (b) ς=12; (c) ς=14; (d) ς=16

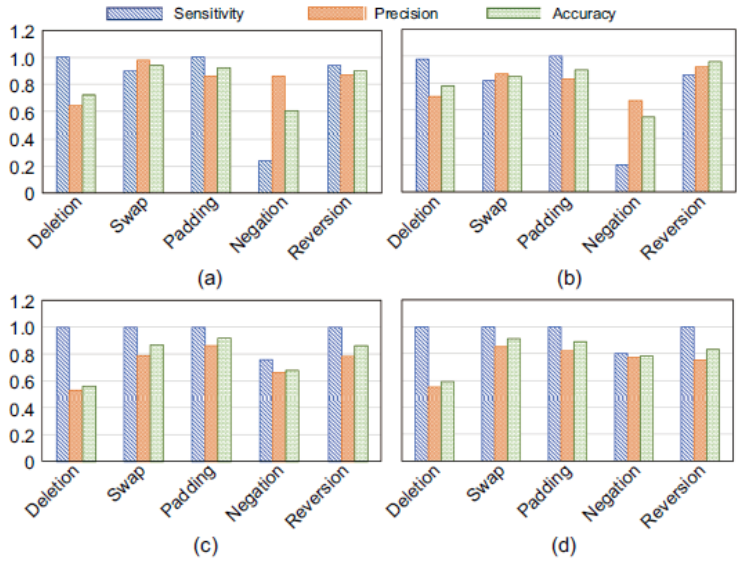


Fig. 10 Robustness of RCDS: (a) ε=0.4, ς=12; (b) ε=0.5, ς=12; (c) ε=0.4, ς=14; (d) ε=0.5, ς=14

Method

RCDS uses both forward validation and backward validation to ensure acceptable efficiency while improving the accuracy of fingerprint identification.

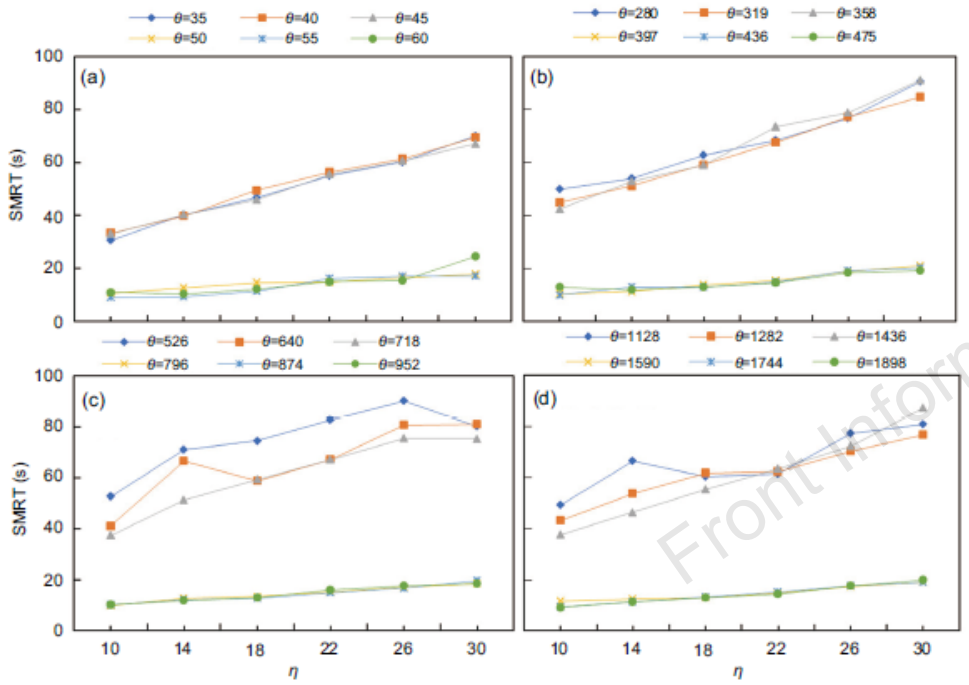


Fig. 11 Symbol mapping generation runtime (SMRT) of RCDS: (a) $|d|$=1 MB; (b) $|d|$=8 MB; (c) $|d|$=16 MB; (d) $|d|$=32 MB

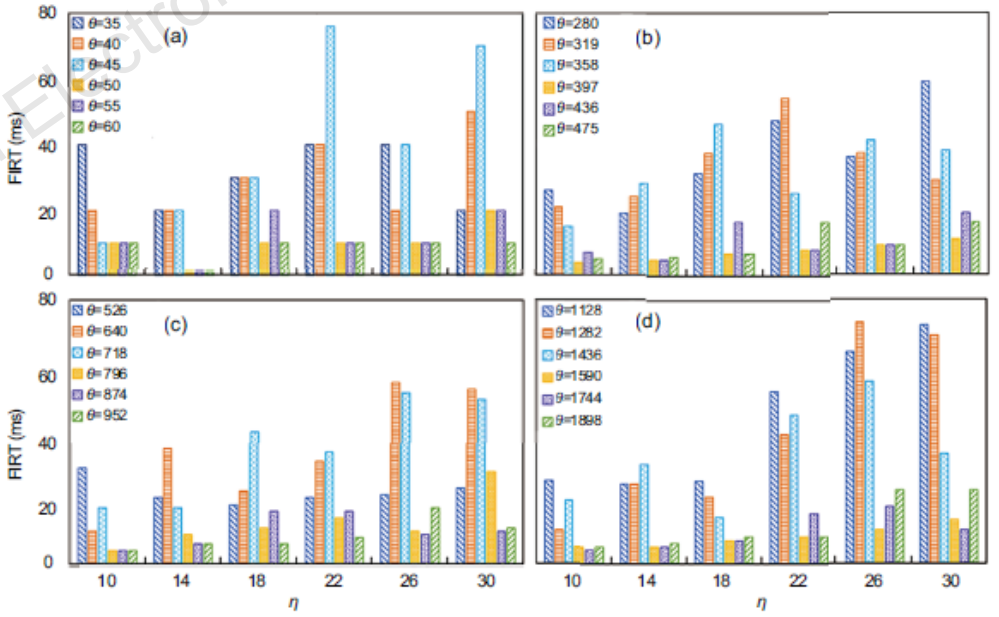


Fig. 12 Fingerprint identification runtime (FIRT) of RCDS: (a) $|d|$=1 MB; (b) $|d|$=8 MB; (c) $|d|$=16 MB; (d) $|d|$=32 MB

Method

The blockchain network performs well with good average delay and throughput when the RCDS model was working.

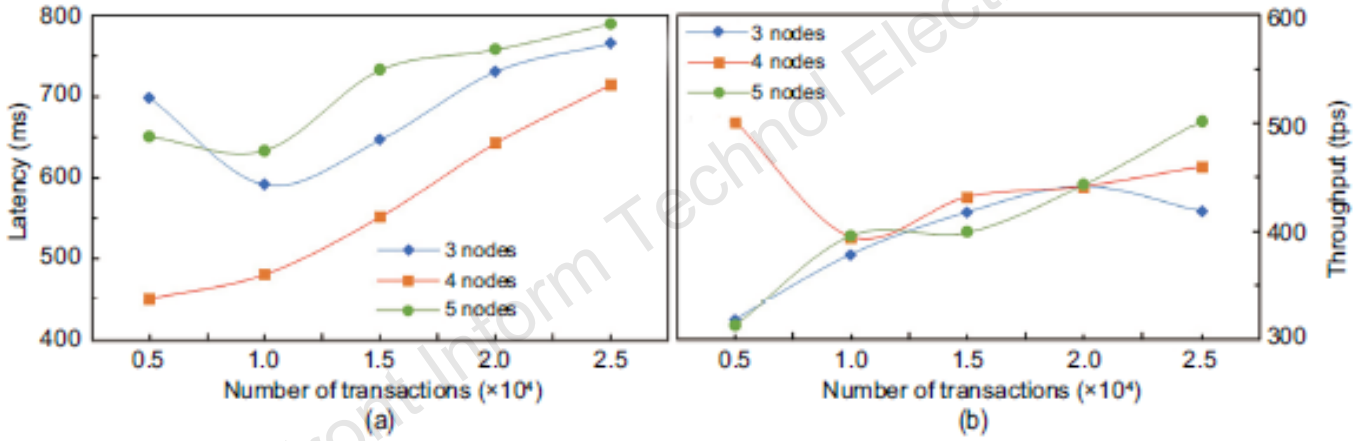
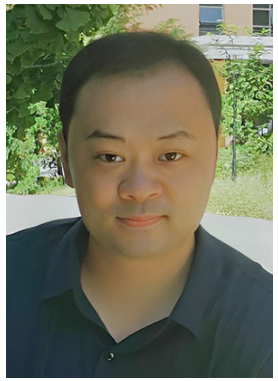


Fig. 13 Comparison of average latency (a) and throughput (b) (tps: transactions per second)

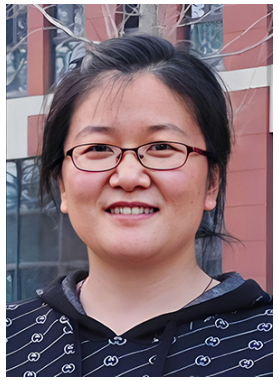
Conclusions

1. By using SMC, RCDS encodes raw data in a non-distortion way, and thus is competent for DRC regardless of the types and volumes of shared data.
2. By employing blockchain, RCDS imposes credible supervision on DRC through the whole network consensus.
3. RCDS combines SMC and blockchain into a systematic mechanism, with which the data access can be fully under control during its sharing processes.
4. Features of RCDS make it possible to launch trusted traitor tracing and access control, better supporting the forensics on the acts of transaction repudiation and data piracy.



Liang WANG is a senior member of CCF and a professional member of ACM/IEEE. He received his PhD degree in computer science and technology from Yanshan University, China. He was previously a visiting scholar at McGill University, Canada. He is now an associate professor at Hebei University. His research interests include blockchain, software security, etc.

Shunjiu HUANG received his BS degree in software engineering from Jiangxi Normal University, Nanchang, China, in 2020, and is currently pursuing his MS degree in electronic and information engineering at Hebei University, Baoding, China. His research interests include blockchain, data sharing, information security, and distributed systems.



Lina ZUO received her MS degree in computer application technology from Hebei University, Baoding, China, in 2008. Her research interests include pattern recognition, graphical information processing, and digital image processing.

Jun LI became a senior member of CCF in 2007. He received his PhD degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, China, in 2006. He is now an professor with the Blockchain Lab., Xiong'an Intelligent City Innovation Federation, China. His research interests include blockchain, data sharing, and distributed digital ID.

Wenyuan LIU received his PhD degree in computer application technology from Harbin Institute of Technology, Harbin, China, in 2000. He is currently a professor and doctoral supervisor with the School of Information Science and Engineering, Yanshan University, Qinhuangdao, China. He is also the CEO of YSUSOFT Information Systems Co., LTD. His research interests include blockchain, data governance, sensor network, information resource planning, information security, and cloud computing.