IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 3, 2018, accepted May 5, 2018, date of publication June 4, 2018, date of current version July 12, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2842245

Information Measurement of Cognitive Communication Systems: The Introduction of Negative Cognitive Information

QIHUI WU[1], **(Senior Member, IEEE), LIZHEN CHEN**[2], **ZHENG WANG**[1], **(Member, IEEE), GUORU DING**[2], **(Senior Member, IEEE), LINYUAN ZHANG**[2], **AND XIAOFEI ZHANG**[1]

[1]Department of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210007, China
[2]College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China

Corresponding author: Zheng Wang (z.wang@ieee.org)

This work was supported in part by the National Natural Science Foundation of China under Grant 61501510 and Grant 61631020 and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150717.

ABSTRACT With rapid developments of artificial intelligence, big data, and data mining, the intelligence-based cognitive communication system has received increasing research attentions, where the traditional communication is strengthened through a real time cognition system over the transmission environment. In this paper, the measurement of cognitive information in cognition systems is investigated. Different from classical Shannon's information theory that only considers information as a probabilistic quantity irrespective of the meaning it conveys, we also take the correctness of cognition into account in the measurement of cognitive information, where the concept of negative cognitive information is introduced for the first time. Specifically, the notion of average cognitive mutual information amount is proposed as a measurement to quantify the cognitive information in average. Then, the concept of cognitive capacity of a given cognition system is defined in terms of the average cognitive mutual information amount, where the maximization or minimization is with respect to the cognitive channel between input and output of a cognition system. Finally, a practical cognitive communication system is presented, where the validity and necessity of the proposed measurement for cognitive information is confirmed.

INDEX TERMS Cognitive communication system, cognitive information, cognitive capacity.

I. INTRODUCTION

The celebrated work that established the discipline of information theory is Claude E. Shannon's landmark paper published in 1948 [1]. In particular, Shannons notion of information was originally developed as means to measure the channel capacity of a communication system, where channel capacity is interpreted as a measure of choice, uncertainty, entropy and lack of knowledge [2]. However, during communication process, Shannon omitted the consideration of semantic meaning carried by the message for the reason that "these semantic aspects of communication are irrelevant to the engineering problem". Therefore, the communication system proposed by Shannon is essentially limited to the communication of data rather than information.

Take the binary pattern recognition system shown in Fig. 1 as an example, since semantic meaning is considered, the cognition performance of systems (a) and (b) are different due to the different cognition meanings of $0 \rightarrow 0$ and $0 \rightarrow 1$. However, according to the measurement of Shannon's information theory, these two cases are equivalent with the same channel capacity $C_a(X;Y) = C_b(X;Y)$ regardless of semantic meanings because Shannon only focuses on the engineering realization of transmission irrespective of the semantic meaning it contains, rendering it inapplicable to semantic cognition systems. In fact, just one year after Shannon introduced his information theory, Weaver, as Shannons co-author of their seminal book published in 1949 [3], formally pointed out the possibility of incorporating semantic information within the overall framework of Shannons theory of communications. Since then, there have been several attempts to define notions of information with meaning [4]–[11]. However, to the best of our knowledge, none of them give the explicit solution for the measurement of the semantic information.

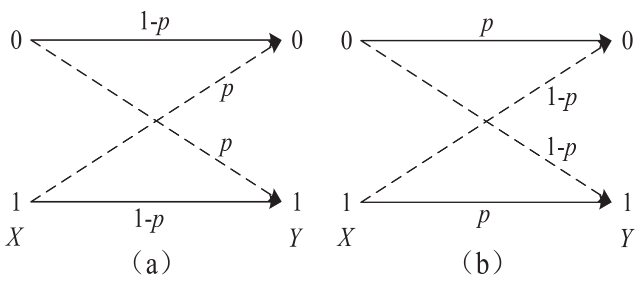


FIGURE 1. Illustration of two binary cognition systems. Cognitions $0 \to 0$ and $1 \to 1$ are depicted by solid curves while cognitions $0 \to 1$ and $1 \to 0$ are plot by dashed curves.

Nowadays, with the upcoming era of artificial intelligence, the demand for information cognition has been far beyond of that for information transmission [12]–[17]. As a matter of fact, many practical experiences have already indicated that the classical Shannon information theory cannot calibrate these advance properly, and the pursuit of the meaning behind the data bits rather than bits itself turns out to be the research direction of the next decades. For example, as shown in Fig. 2, to pursuit more reliable communication, cognition-based intelligence has been adopted to traditional communications, where cognition is performed with respect to the transmission environment. However, one fact has been recognized that the key assumption from Shannon that "semantics is not relevant" no longer holds in the field of information cognition, which implies the theoretic foundation from Shannon may be not well suited for cognition systems [18]–[21].

In this paper, the measurement of cognitive information is studied in full details. To the best of our knowledge, this is the first time the measurement of cognitive information is investigated. For the consideration of semantic aspect, the correctness of cognition is taken into account. Based on it, the concept of average cognitive mutual information is proposed in Section II to quantify the cognitive information in average. Compared to the non-negative average mutual information in communications, it can be either positive or negative, thus leading to beneficial cognition and harmful cognition, respectively. Subsequently, from this raises the definition of cognitive capacity in Section III, which systematically evaluates the cognition ability of a cognitive system. By adjusting the cognition channels, the positive or negative cognition capacity is achieved. In Section IV, two examples of cognitive information in cognitive communications are illustrated. At the end, Section V concludes the paper. To summarize, we contribute to the basic measurement of cognitive information in cognitive communication systems from the following two-fold aspects:

1) Based on the correctness of cognition, propose the concept of average cognitive mutual information for the measurement of cognitive information.
2) Based on the correct and wrong cognitive information amount, define the cognitive capacity to evaluate the maximal cognition ability of a cognitive system.

II. MEASUREMENT OF COGNITIVE INFORMATION AMOUNT

The way of quantifying the information amount by probability can be originally traced back to the work of Hartley [22], and was further generalized by Shannon. Specifically, the information amount of a specific message or event x_i with probability $p(x_i)$ is quantified as

$$I(x_i) = -\log p(x_i). \tag{1}$$

Then, given $I(x_i)$, the information amount of x_i due to the knowledge of y_j becomes

$$I(x_i|y_j) = -\log p(x_i|y_j), \tag{2}$$

and the mutual information $I(x_i; y_j)$ between x_i and y_j is written by

$$\begin{aligned} I(x_i; y_j) &= I(x_i) - I(x_i|y_j) \\ &= \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \end{aligned} \tag{3}$$

Because Shannon only considers information as a probabilistic quantity regardless of its semantic meaning, $I(x_i; y_j)$ is actually calculated by the variation of information amount of x_i given y_j. Therefore, $I(x_i; y_j)$ can be either positive or negative, depending on the ratio between posterior probability and prior probability, i.e., $\frac{p(x_i|y_j)}{p(x_i)}$ or $\frac{p(y_j|x_i)}{p(y_j)}$. Unless stated otherwise, the units of mutual information throughput the context are bits with log base 2. More specifically, when $I(x_i; y_j) \geq 0$, it means the knowledge of y_j is helpful to confirm x_i, otherwise confusion will be introduced by y_j to determine x_i.

Motivated by the semantic demand from cognition field, here we introduce the correctness of cognition into the measurement of information by semantic meanings. To start with, we firstly define the correct and wrong cognition as follows,

$$\text{correct cognition}: \ x_i \to y_j, \quad \text{for } i = j; \tag{4}$$

$$\text{wrong cognition}: \ x_i \to y_j, \quad \text{for } i \neq j. \tag{5}$$

Therefore, the cognition systems can also be simplified as the model of communications, except using correct and wrong cognition channels instead of communication channels. Here we only consider the case that the state space of X and Y are one-to-one correspondence. We admit that the case beyond one-to-one correspondence does exist, which is outside the scope of this survey and will be one of our research work in future.

According to correctness of cognition, the measurement of information by Shannon is refined by semantic meanings, which consists of the following two basic steps:

1) *Classify the variation of information amount within cognition systems as correct cognitive information amount and wrong cognitive information amount,*

$$I_{i=j}(x_i; y_j) = [I(x_i) - I(x_i|y_j)]_+, \tag{6}$$

$$I_{i\neq j}(x_i; y_j) = [I(x_i) - I(x_i|y_j)]_- \tag{7}$$

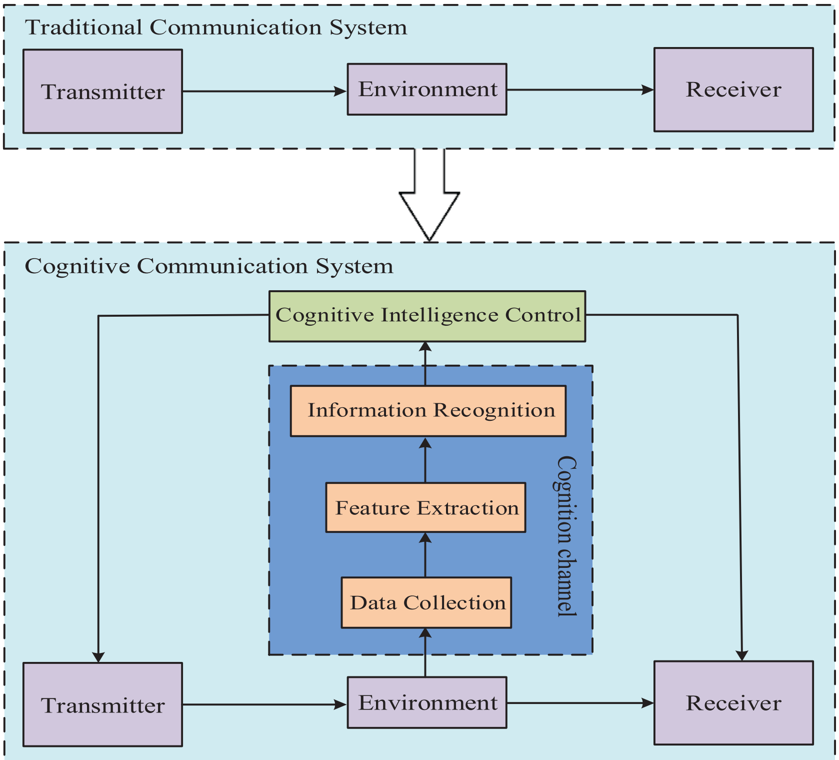


FIGURE 2. **The transition of communication paradigms.**

where $[\cdot]_+ = +(\cdot)$ *and* $[\cdot]_- = -(\cdot)$ *denote the signs of correct cognition and wrong cognition.*

2) *Unify the correct cognitive information amount and wrong cognitive information amount with cognitive information amount by*

$$[I(x_i) - I(x_i|y_j)]_+ = I(x_i) - I(x_i|y_j); \tag{8}$$
$$[I(x_i) - I(x_i|y_j)]_- = -(I(x_i) - I(x_i|y_j)). \tag{9}$$

To summarize, the cognitive mutual information $I_c(x_i; y_j)$ is defined by semantic meanings as

$$I_c(x_i; y_j) = \begin{cases} I(x_i) - I(x_i|y_j) & \text{for } i = j, \\ -(I(x_i) - I(x_i|y_j)) & \text{for } i \neq j. \end{cases} \tag{10}$$

By doing this, the traditional mutual information provided by Shannon is converted to the cognitive mutual information for the sake of semantic consideration. As shown in Fig. 3, by Shannon's measurement, correct and wrong cognition information amounts are scaled according to their exact values regardless of the semantic meaning. As a comparison, cognitive information amount takes the correctness of cognition into account. Specifically, given the self-information amount $I(x_i)$, the more information decrement from $I(x_i)$ to $I(x_i|y_j)$ by correct cognition with $i = j$, the more beneficial of this cognition, and vice versa. To be more particular, as for cognition with respect to "0", the decrement of information from $I(0)$ to $I(0|0)$ by correct cognition is beneficial and the decrement of information between $I(0)$ and $I(0|1)$ by wrong cognition becomes harmful.

III. MEASUREMENT OF COGNITIVE CAPACITY

Based on the mutual information $I(x_i, y_j)$, the average mutual information between two random variables X and Y is further defined by Shannon as

$$\begin{aligned} I(X; Y) &= \sum_{x_i \in \mathcal{X}} \sum_{y_j \in \mathcal{Y}} p(x_i, y_j) I(x_i; y_j) \\ &= \sum_{x_i \in \mathcal{X}} \sum_{y_j \in \mathcal{Y}} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \geq 0. \end{aligned} \tag{11}$$

Because average mutual information $I(X; Y)$ measures entropy (i.e., uncertainty) reduction of X due to the knowledge of Y, it naturally can be applied to evaluate the commonness between X and Y. By simply extending X and Y to communications as information source and destination respectively, the channel capacity that describes the transmission ability of a general communication system is defined by

$$C = \max_{p(X)} I(X; Y), \tag{12}$$

where the maximization is with respect to the design of the input distribution $p(X)$. However, during the communication process, the consideration of semantic meaning is ignored, making the two cases (a) and (b) in Fig. 1 have exactly the same channel capacity.

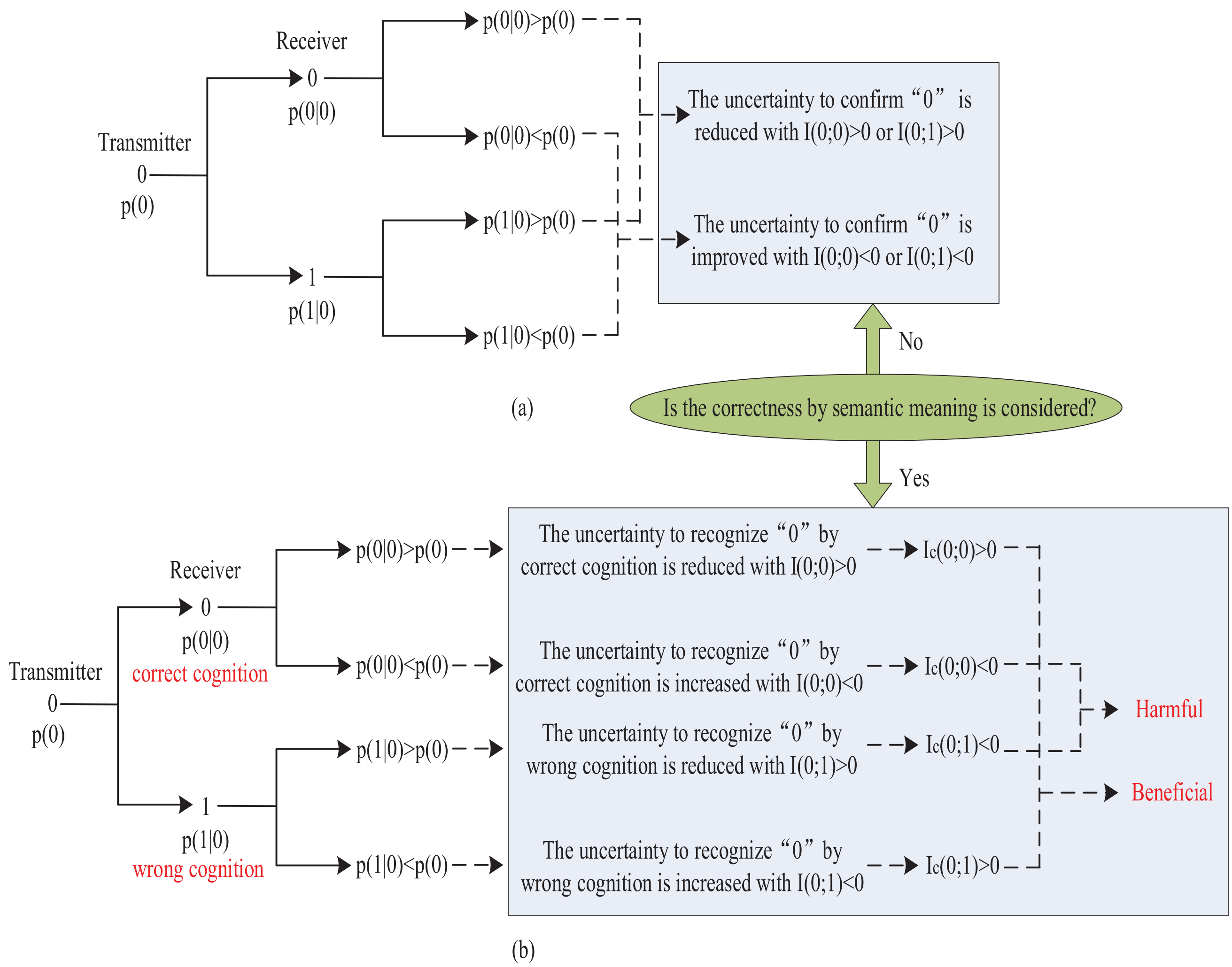


FIGURE 3. The illustrations of Shannon's mutual information (a) and cognitive mutual information (b) with respect to "0".

On the other hand, different from Shannon, the correctness of semantic meaning is urgent to be considered in cognitive communications. Therefore, based on the cognitive mutual information shown in (10), the correct and wrong cognitive information amount are defined respectively as

$$I_{c+}(X;Y) = \sum_{i=j} p(x_i, y_j) I_{i=j}(x_i; y_j), \tag{13}$$

$$I_{c-}(X;Y) = \sum_{i\neq j} p(x_i, y_j) I_{i\neq j}(x_i; y_j); \tag{14}$$

while the average cognitive mutual information is defined as the summation of average correct and wrong cognitive information amount in average:

$$I_c(X;Y) = P_{c+} \cdot I_{c+}(X;Y) + P_{c-} \cdot I_{c-}(X;Y), \tag{15}$$

where $P_{c+} = \sum_{i=j} p(x_i, y_j)$ and $P_{c-} = \sum_{i\neq j} p(x_i, y_j)$. Furthermore, from (15), it is easy to verify that the average cognitive mutual information $I_c(X;Y)$ is bounded as

$$-H(X) \leq I_c(X;Y) \leq H(X), \tag{16}$$

where $H(X) = -\sum_{i\in\mathcal{X}} p(x_i) \log p(x_i)$ is the entropy of X defined by Shannon. Therefore, a salient feature of cognitive mutual information is that it could be negative in some cases of interest.

As shown clearly in Fig. 4, in a binary symmetric channel, the proposed average cognitive mutual information $I_c(X;Y)$ behaves different from the mutual information $I(X;Y)$ given by Shannon. More specifically, the cognition system tends to be beneficial for $I_c(X;Y) > 0$ while harmful understanding will dominate the cognition if $I_c(X;Y) < 0$. As for the case $I_c(X;Y) = 0$, it implies the positive cognition is as the same as the negative one, resulting in an invalid cognition. Therefore, it can be used to evaluate how faithful a cognition system is.

Next, we define the positive cognitive capacity as the maximum of average cognitive mutual information to evaluate the positive cognitive ability of a cognition system:

$$C_{\text{cognitive}+} = \max_{p(y_j|x_i)} I_c(X;Y), \tag{17}$$

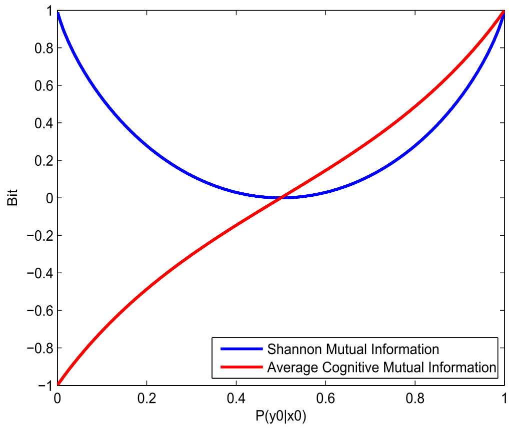


FIGURE 4. Illustration of Shannon mutual information and average cognitive mutual information under a binary symmetric channel with $p(y_0|x_0) = p(y_1|x_1) = 1 - p(y_1|x_0) = 1 - p(y_0|x_1)$ and $p(x_0) = p(x_1) = \frac{1}{2}$.

where $p(y_j|x_i)$ denotes the cognitive transition probability for a specific cognition channel from x_i to y_j. In theory, positive cognitive capacity serves as a tight upper on the rate at which information can be reliably recognized over a cognition process. Apart from channel capacity shown in (12), the maximization of the positive cognitive capacity is with respect to the cognition channels $p(y_j|x_i)$, and it can be realized by improving the quality of the cognition channels. For example, if $p(y_j|x_i) = 1$ for all $i = j$, it corresponds to a perfect cognition without wrong cognition, i.e., $p(y_j|x_i) = 0$ for $i \neq j$, resulting in

$$C_{\text{cognitive}+} = H(X). \quad (18)$$

As for optimizing $p(y_j|x_i)$, the reasonable feature selection is the key to exploit [23], [24], followed by the steps of data collection, feature extraction, semantic recognition and so on, which will be a main direction of our future work. Consequently, with the increment of $p(y_j|x_i)$ for $i = j$, the average cognitive mutual information will gradually approach its positive capacity.

On the other hand, cognitive information can also be used in military for security attack. From this perspective, how to lower the average cognitive mutual information $I_c(X; Y)$ in hostile cognition systems becomes the key of interest, which is implemented by disguising the cognition channels for misinformation or disinformation. Therefore, the negative cognition capacity is defined as

$$C_{\text{cognitive}-} = \min_{p(y_j|x_i)} I_c(X; Y) = -H(X), \quad (19)$$

which serves as a lower bound on the rate at which the information is misrecognized over a cognition process. In the scenario of military, the notion of negative cognition capacity plays a crucial role by misleading and destroying the cognition systems on the other side. More specifically, by deliberately deteriorating the recognition channels, the average cognitive mutual information will achieve this negative capacity when $p(y_j|x_i) = 1$ for $i \neq j$ and $p(y_j|x_i) = 0$ for $i = j$.

IV. COGNITIVE INFORMATION IN COGNITIVE RADIO NETWORK

A cognitive radio network enables the secondary user to utilize the spectrum not currently being by the primary user, known as a spectrum hole, after performing sensing on the spectrum [25]. In order to improve the accuracy of spectrum sensing, cooperative spectrum sensing (CSS) is applied, where sensing results of multiple spectrum sensors are reported to the secondary user for a better spectrum utilization. However, the cognition process may be deteriorated in a hostile environment, where some sensors are malicious and falsify the sensing results to mislead the secondary user [26].

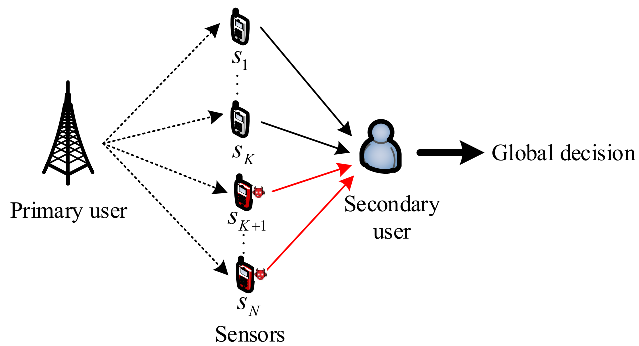


FIGURE 5. Illustrations of cooperative spectrum sensing in a hostile cognitive radio network.

Just as shown in Fig. 5, there are N sensing nodes in the CSS network, including K honest nodes, i.e., $s_1, s_2, \ldots, s_K$, and M malicious nodes with $K + M = N$, i.e., $s_{K+1}, s_{K+2}, \ldots, s_N$. Specifically, as for the honest sensing node, let its correct sensing probabilities $p_h(y_0|x_0)$ and $p_h(y_1|x_1)$ are represented by p. Accordingly, the wrong sensing probabilities $p_h(y_1|x_0)$ and $p_h(y_0|x_1)$ are $1 - p$. Meanwhile, as for the malicious sensing node reversing its sensing results with an attack probability p_a, its correct and wrong sensing probabilities are disturbed, that is

$$p_m(y_0|x_0) = p_m(y_1|x_1) = p \cdot (1 - p_a) + (1 - p) \cdot p_a, \quad (20)$$

$$p_m(y_0|x_1) = p_m(y_1|x_0) = (1 - p) \cdot (1 - p_a) + p \cdot p_a. \quad (21)$$

Overall, given the sensing results from N sensing nodes, the secondary user makes data fusion and global decision. Here, the majority rule is exploited to make decisions, i.e., if and only if there are no fewer than L sensors reporting the absence of the primary user, the global decision is y_1, and otherwise, y_0. Hence, the transition probabilities of cognition channels in the CSS network can be derived as

$$p_{css}(y_0|x_0) = p_{css}(y_1|x_1) = \sum_{l=L}^{N} \sum_{j=a}^{b} f(j; M, p_m(y_0|x_0)) \cdot f(l - j; K, p_h(y_0|x_0)), \quad (22)$$

where $a = \max(0, l - K)$, $b = \min(l, M)$, $f(v, w, e) = [w\ v]^T \cdot e^v \cdot (1 - e)^{w-v}$. Further, we have

$$p_{css}(y_0|x_1) = p_{css}(y_0|x_1) = 1 - p_{css}(y_0|x_0). \quad (23)$$

Then, from (22) and (23), the average cognitive mutual information of this CSS network system can be calculated based on (15). Just as depicted in Fig. 6, the Shannon mutual information $I(X;Y)$ and average cognitive mutual information $I_c(X;Y)$ show distinct trends over the attack percentage which is the ratio of malicious nodes to all nodes, i.e., M/N. Specifically, as the attack percentage increases, the performance of CSS deteriorate and the average cognitive mutual information decreases, which is easy to understanding. In contrast, the Shannon mutual information increases when the attack percentage is over certain level, which is not consistent with the cognition performance. Furthermore, when the attack probability p_a decreases from 1 to 0.8, the cognition performance becomes better, and $I_c(X;Y)$ increases. However, the Shannon mutual information $I(X;Y)$ shows heterogeneous changes, that is, if the attack percentage is lower than 0.55, $I(X;Y)$ increases, and otherwise, $I(X;Y)$ decreases. Hence, compared to the Shannon mutual information, the proposed average cognitive mutual information well formulates the cognition performance.

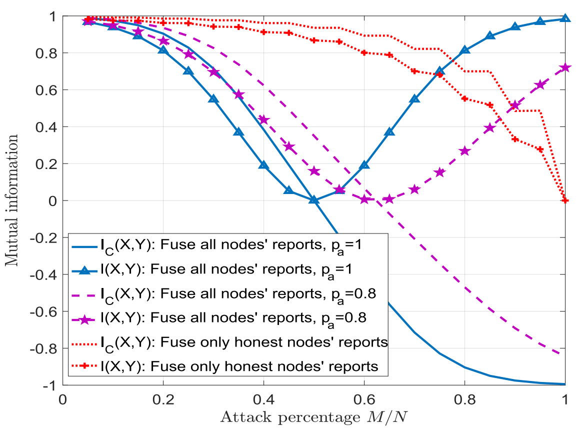


FIGURE 6. Illustration of Shannon mutual information and average cognitive mutual information for cooperative spectrum sensing in a hostile cognitive radio network, where $p = 0.8$, $N = 20$, $L = 10$.

On the other hand, some classic defense strategies are introduced to eliminate negative effects of data falsification and the defense performance is evaluated via the average cognitive mutual information in Fig. 7. Here, three strategies are considered, i.e.,

- Global filtering: global decisions are used as a reference to make comparison with nodes' reports and nodes with high inconsistency are identified as malicious and forbidden to participate into data fusion [27].
- Trusted-node assisting: an honest node is prior known by the SU and its reports are used to compared with others'. Nodes with high inconsistency are identified as malicious and forbidden to participate into data fusion [28].
- Optimal fusion with estimated parameters: every node's real performance is evaluated by a reliable reference, based on which an optimal likelihood ratio test is made [29].

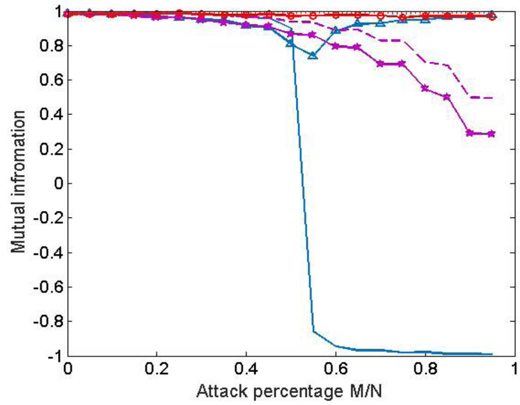


FIGURE 7. Performance of difference defense strategies under various attack percentage, where the attack probability is 1.

As global decisions are sensitive to the attack percentage, the first strategys performance decreases dramatically with the attack percentage. Specially, when the attack percentage is over 0.5, honest nodes are identified as malicious and malicious ones are identified as honest. It can be well reflected by $I_c(X;Y)$, but the Shannon mutual information doesn't correctly reflect the trend. Although the trust node's performance is immune to the attack and the strategy, trusted-node assisting, can well identify the attributes of nodes, the number of nodes participating in data fusion decreases with the attack percentage, and the performance deteriorates. Differently, in the third scheme, as the real performance of nodes are estimated, the reports of attackers are reversed via the likelihood ratio test and their cognitive mutual information turn from negative to positive. Hence, the global performance is good and the mutual information is large.

V. CONCLUSION AND FUTURE WORK

Following the footstep of Shannon and Weaver, in this paper, we try to define the measurement of semantic information in cognition communication systems. Our results extend the classical Shannon's information theory by attempting to characterize the semantic meaning conveyed in the information. More specifically, the correctness of cognition information is introduced to measure the amount of cognitive information. As shown in Fig. 4, one elegant result is that the capacity of cognitive information, in terms of average cognitive mutual information, monotonously increases with the probabilities of correction cognition (i.e., quality of cognitive channels), which can serve as a fundamental performance metric for multi-disciplinary applications, such as classification in pattern recognition and machine learning [12], environment sensing in internet of things [19], signal detection in cognitive communications [18], to name just a few.

Semantic information is a good complement and extension to Shannonian information, which will be a fruitful research direction in the future. Therefore, we firmly believe that our current work have only touched one tip of an iceberg, and there are lots of open questions to be addressed. For example, through the context, the semantic set $\mathcal{X} = \{x_1, \ldots, x_n\}$ or $\mathcal{Y} = \{y_1, \ldots, y_n\}$ we investigated for random variables X or Y is separable, namely, $x_i \cap x_j = \emptyset$ and $y_i \cap y_j = \emptyset$. However, in practice, the elements in the semantic set tend to be partially overlapped due to the ambiguity contained in the semantic text. Besides, the feature selection also has an impact on the semantic ambiguity of the cognition, making the reasonable feature selection a crucial problem in cognition systems. In addition, Fig. 4 only illustrates an example of binary cognition, where complicated cognition systems also can be equipped with the proposed cognitive information. Therefore, we hope this article will stimulate much more research interest.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[2] S. Verdú, "Fifty years of Shannon theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2057–2078, Oct. 1998.

[3] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, IL, USA: Univ. of Illinois Press, 1949.

[4] T. T. Rogers and J. L. McClelland, *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA, USA: MIT Press, 2004.

[5] H. Haken and J. Portugali, *Information Adaptation: The Interplay Between Shannon Information and Semantic Information in Cognition*. Springer, 2015.

[6] L. Floridi, "Semantic conceptions of information," Stanford Encyclopedia of Philosophy, Tech. Rep., 2011.

[7] J. Bao *et al.*, "Towards a theory of semantic communication," in *Proc. 1st IEEE Int. Workshop Netw. Sci.*, Jun. 2011, pp. 110–117.

[8] R. B. Wells, "Weaver's model of communication and its implications," Univ. Idaho Reas. Expertise, Moscow, ID, USA, Tech. Rep., 2011, vol. 17, pp. 1–20. [Online]. Available: https://vivo.nkn.uidaho.edu/vivo/display/n130915

[9] D. H. Li, "The semantic aspect of communication theory and accountancy," *J. Accounting Res.*, vol. 1, no. 1, pp. 102–107, 1963.

[10] V. Bojilov, "Formal theory of semantic and pragmatic information—A technocratic approach," *Int. J. Inf. Theories Appl.*, vol. 22, no. 4, pp. 356–397, 2015.

[11] S. Feng, Q. Wu, J. Wang, Y. Xu, and G. Ding, "An introduction of cognition information: From form aspect to semantic aspect," in *Proc. 22nd Wireless Opt. Commun. Conf.*, 2013, pp. 481–485.

[12] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[13] R. Michalski, J. Carbonell, and T. Mitchell, *Machine Learning: An Artificial Intelligence Approach*. Berlin, Germany: Springer, 2013.

[14] M. J. Berry and G. S. Linoff, *Data Mining Techniques: For Marketing, Sales, and Customer Support*. Hoboken, NJ, USA: Wiley, 1997.

[15] J. Andrews *et al.*, "Rethinking information theory for mobile ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 12, pp. 94–101, Dec. 2008.

[16] T. L. Bauer, "Information and meaning: Revisiting Shannon's theory of communication and extending it to address today's technical problems," Sandia Nat. Lab., Albuquerque, NM, USA, SANDIA Rep. SAND2009-8168, 2009, pp. 1–40.

[17] A. Sheth, "Internet of Things to smart IoT through semantic, cognitive, and perceptual computing," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 108–112, Mar. 2016.

[18] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[19] Q. Wu *et al.*, "Cognitive Internet of Things: A new paradigm beyond connection," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 129–143, Apr. 2014.

[20] C. Thorton, "A new way of linking information theory with cognitive science," in *Proc. Annu. Meeting Cogn. Sci. Soc.*, vol. 35, 2013, pp. 3545–3550.

[21] A. Nematzadeh, A. Fazly, and S. Stevenson, "A cognitive model of semantic network learning," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 244–254.

[22] R. V. L. Hartley, "Transmission of information," *Bell Labs Tech. J.*, vol. 7, no. 3, pp. 535–563, 1928.

[23] G. Brown, "A new perspective for information theoretic feature selection," in *Proc. Artif. Intell. Statist.*, 2009, pp. 49–56.

[24] M. Last, A. Kandel, and O. Maimon, "Information-theoretic algorithm for feature selection," *Pattern Recognit. Lett.*, vol. 22, nos. 6–7, pp. 799–811, 2001.

[25] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 126–136, Jul. 2013.

[26] L. Zhang, G. Ding, Q. Wu, Y. Zou, Z. Han, and J. Wang, "Byzantine attack and defense in cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1342–1363, 3rd Quart., 2015.

[27] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 774–786, Feb. 2011.

[28] K. Zeng, P. Paweczak, and D. Čabrić, "Reputation-based cooperative spectrum sensing with trusted nodes assistance," *IEEE Commun. Lett.*, vol. 14, no. 3, pp. 226–228, Mar. 2010.

[29] L. Zhang, G. Ding, Q. Wu, and F. Song, "Defending against Byzantine attack in cooperative spectrum sensing: Defense reference and performance analysis," *IEEE Access*, vol. 4, pp. 4011–4024, 2016.

QIHUI WU (SM'13) received the B.S. degree in communications engineering, the M.S. degree, and the Ph.D. degree in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1994, 1997, and 2000, respectively. From 2003 to 2005, he was a Post-Doctoral Research Associate with Southeast University, Nanjing, China. From 2005 to 2007, he was an Associate Professor with the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, where he served as a Full Professor from 2008 to 2016. In2011, he was an Advanced Visiting Scholar with the Stevens Institute of Technology, Hoboken, NJ, USA. Since 2016, he has been a Full Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests span the areas of wireless communications and statistical signal processing, with emphasis on system design of software defined radio, cognitive radio, and smart radio.

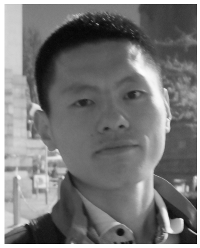
LIZHEN CHEN received the B.S. and M.S. degrees from Qufu Normal University, Shandong, China, in 2002 and 2009, respectively. She is currently pursuing the Ph.D. degree in communication engineering with Army Engineering University of PLA, China. She was a Lecturer with Shandong Agricultural University. Her research interests are in the areas of information theory and cognitive radio.

ZHENG WANG received the B.S. degree in electronic and information engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2009, the M.S. degree in communications from the Department of Electrical and Electronic Engineering, University of Manchester, Manchester, U.K., in 2010, and the Ph.D. degree in communication engineering from Imperial College London, UK, in 2015.

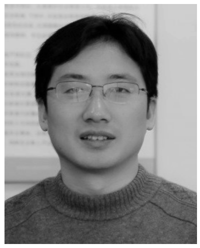
From 2015 to 2016, he served as a Research Associate with Imperial College London, U.K. and from 2016 to 2017, he was a Senior Engineer with the Radio Access Network R&D Division, Huawei Technologies Co. He is currently an Assistant Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include lattice methods for wireless communications, cognitive radio, and physical layer security.

GUORU DING (S'10–M'14–SM'16) received the B.S. degree (Hons.) in electrical engineering from Xidian University, Xi'an, China, in 2008, and the Ph.D. degree (Hons.) in communications and information systems from the College of Communications Engineering, Nanjing, China, in 2014. Since 2014, he has been an Assistant Professor with the College of Communications Engineering and a Research Fellow with the National High Frequency Communications Research Center of China. Since 2015, he has been a Post-Doctoral Research Associate with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. His research interests include cognitive radio networks, massive MIMO, machine learning, and big data analytics over wireless networks.

He has acted as Technical Program Committees member for a number of international conferences, including the IEEE Global Communications Conference, the IEEE International Conference on Communications, and the IEEE Vehicular Technology Conference. He is a Voting Member of the IEEE 1900.6 Standard Association Working Group. He was a recipient of the Best Paper Awards from EAI MLICOM 2016, IEEE VTC 2014-Fall, and IEEE WCSP 2009. He was a recipient of the Alexander von Humboldt Fellowship in 2017 and the Excellent Doctoral Thesis Award of the China Institute of Communications in 2016. He has served as a Guest Editor of the IEEE Journal on Selected Areas in Communications (Special issue on spectrum sharing and aggregation in future wireless networks). He is now an Associate Editor of the Journal of Communications and Information Networks, the KSII *Transactions on Internet and Information Systems* and the *AEU-International Journal of Electronics and Communications*.



LINYUAN ZHANG received the B.S. degree (Hons.) in electronic engineering from Inner Mongolia University, Hohhot, China, in 2012. He is currently pursuing the Ph.D. degree in communications and information system with the College of Communications Engineering, Army Engineering University of PLA. His research interests are wireless security and statistical learning.



XIAOFEI ZHANG received the M.S. degree from Wuhan University, Wuhan, China, in 2001, and the Ph.D. degree in communication and information systems from the Nanjing University of Aeronautics and Astronautics in 2005. He is currently a Full Professor with the Electronic Engineering Department, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research is focused on array signal processing and communication signal processing.

Dr. Zhang serves on the Technical Program Committees of the IEEE 2010 International Conference on Wireless Communications and Signal Processing, the IEEE 2011 International Conference on Wireless Communications and Signal Processing, ssme2010, and 2011 International Workshop on Computation Theory and Information Technology. He serves as an Editor of the *International Journal of Digital Content Technology and its Applications*, the *International Journal of Technology and Applied Science*, the *Journal of Communications and Information Sciences*, the *Scientific Journal of Microelectronics and International Journal of Information Engineering*.

He also serves regularly as a Peer Reviewer for the IEEE Transactions Wireless Communication, the IEEE Transactions on Vehicular Technology, *EURASIP Journal on Advances in Signal Processing*, the IEEE Communication Letters, *Signal Processing*, the *International Journal of Electronics*, the *International Journal of Communication Systems*, and *Wireless Communications and Mobile Computing*.