Understand-Before-Talk (UBT): A Semantic Communication Approach to 6G Networks

Shiva Raj Pokhrel, Senior Member, IEEE, and Jinho Choi, Fellow, IEEE

Abstract—In Shannon theory, semantic aspects of communication were identified but considered irrelevant to the technical communication problems. Semantic communication (SC) techniques have recently attracted renewed research interests in 6th generation (6G) wireless, because they have the capability to support an efficient interpretation of the significance and meaning intended by a sender (or accomplishment of the goal) when dealing with multi-modal data such as videos, images, audio, text messages, and so on, which would be the case for various applications such as intelligent transportation systems where each autonomous vehicle needs to deal with real-time videos and data from a number of sensors including radars. To this end, most of the emerging SC works focus on specific data types and employ sophisticated machine learning models including deep learning and neural networks. However, they could be impractical for multi-modal data possibly within a real-time constraint, relative to the purpose of the communication. A notable difficulty of existing SC frameworks lies in handling the discrete constraints imposed on the pursued semantic coding and its interaction with the independent knowledge-base, which makes reliable semantic extraction extremely challenging. Therefore, we develop a new hashing-based semantic extraction approach to SC framework, where our learning objective is to generate one time signatures (hash codes) using supervised learning for low latency, secure and efficient management of the SC dynamics. We first evaluate the proposed semantic extraction framework over large image data sets, extend it with domain adaptive hashing and then demonstrate the effectiveness of “semantics signature” in bulk transmission and multi-modal data.

Index Terms—Semantic communication, semantic extraction, domain adaptation, hashing, semantics signature, fast supervised learning, multimodal dataset, stability and convergence.

I. INTRODUCTION

NIKOLA Tesla once said: “When wireless is perfectly applied the whole earth will be converted into a huge brain, which in fact it is, all things being particles of a real and rhythmic whole,” as Tesla envisioned today’s mobile phones and future sixth-generation (6G) networks back in 1926.

In Shannon theory, semantic aspects of communication were considered irrelevant to the technical communication problems [1]. From earlier [1] and recent works [2], [3], [4], [5], [6] in the literature, and motivated by the above Tesla’s quote, we found that 6G design innovations can get valuable suggestions by cautiously observing how our brain’s processes environmental inputs gathered by our senses. Our brain learns, in substance, from prior acts (and also from our historically collected understandings and culture) and executes complicated reasoning by using sustainable energy in a short time. Such semantic aspects (messages have meaning, and they refer to or are correlated according to some systems with certain physical or conceptual entities [1]), of communication become significant for 6G [6].

In 6G networks, inheriting these excellent human brain mechanisms given by nature, should include semanticization and effectiveness in communications as the fundamental components for imminent innovations [7]. As shown in Fig. 1, such a mechanism allows us to distill the data that are strictly relevant for conveying (semantic) information from sender to receiver and focus on to clearly establish the purpose of communication. For example, data-aided sensing (DAS) proposed in [8] learns from prior acts and can be a valuable technique to considerably minimize the quantity of data to be transmitted (by decreasing irrelevant data transfer and using historically acquired experience).

Semantic communication, as illustrated in Fig. 1, roots back to the three levels founded by Shannon and Weaver:

- **Technical (Level A):** How correctly can the bits/packets be transmitted from the sender to the receiver?
- **Semantics (Level B):** How accurately do the bits/packets carry the required meaning from the sender to the receiver?
Effectiveness (Level C): How well does the meaning transmitted (conveyed) influence the intended course of action at the receiver?

At level A, semantic aspects of communication are assumed to be irrelevant to the technical engineering problems. Shannon noted “Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem” [1]. This is one of the fundamental assumptions of Shannon’s theory. The relaxation of this fundamental assumption for communication between a sender and a receiver can jointly consider all three levels of problem, which can be referred to as semantic communication (SC). In addition, SC can borrow valuable insights from Sperber and Wilson’s inferential model in communication [9]–‘human cognition has a goal’.

A. Motivation

Conventional communication systems’ actual transmission rates are getting close to the Shannon limit, and the remaining spectrum resources are being increasingly scarce [7], [10]. The conversion of the transmission-before-understanding communication paradigm to the understanding-before-transmission paradigm will be the key to SC’s success in addressing bandwidth bottlenecks and spectrum scarcity. This is referred to as Understand-Before-Talk (UBT) in this paper. By doing this, semantic extraction can be included into the communication model to create SC for 6G [11], which only permits the transmission of information that is relevant to the recipient and hides or reduces redundant data, relieving bandwidth pressure and improving privacy protection [7], [12]. However, accuracy of semantic extraction has been highly challenging and non-trivial task in SC engineering.

One of the fundamental simplifying assumptions of Shannon’s theory is that the information or the semanticized content is inherent in the transmitted message bits and that it is self-sufficient to retrieve the semantic information, once acquired by the receiver. Beyond the scope of Shannon’s theory, a tightly coupled synchronization between sender and receiver is essential for successfully retrieving semantics at both ends. In contrast to the recent developments [2], [7], [13], [14], [15], [16], [17], detailed later in Section I-B and II-A, our objective is to develop an accurate semantic extraction approach for SC to facilitate joint optimization (information gathering, information dissemination, and decision-making) over 6G networks.

We design a hashing-based semantic extraction framework for generating signatures enabling a low-delay generic SC approach applicable across all message types. Different from other works, we are motivated by the human communication’s realities, which demonstrate that what we send materially is entirely insufficient in interpreting the actual allusions that the sender meant to mean by transmitting the message. Main components essential for indicating or acting in semantic sense, are: i) surroundings and situations, ii) knowledge sharing, iii) prior communications held (used in [8]), iv) using empathy by sender/receiver, etc.

Our preliminary moves towards this end are the studies towards employing federated learning [18] of connected vehicles for knowledge sharing [5] and automation [19], probabilistic inference based extraction [20] and the data-aided sensing [8] (by taking into account the prior communications held). Indeed, it is hard to ascertain that the information is contained just in the message sent, rather than which, we should understand messages to be a key to information access (a more sophisticated process) and this could be the core of 6G communication [5], [21].

B. Impeding Challenges in SC

Despite the fact that SC has been viewed by many new approaches as a robust 6G technique, there are still several issues to be addressed. We provide a synopsis of three typical challenges in SC and the relevant technologies and methodologies that are available for various uses.

a) Accuracy of Semantic Extraction: Semantic extraction [37] has been quite successful in different contexts, such as computer vision (e.g., semantic segmentation), semantic computing and semantic web [38] (similar to knowledge graph) for recommendation systems to facilitate intelligent and high quality user experience [39]. SC presents challenges for accurate semantic extraction, particularly for the communication entities whose context may evolve individually, because all entities need to be strongly aligned semantic interpretation and representation. Additionally, data communication in 6G has a high accuracy and a strong time-sensitive requirements (so called extreme URLLC [40]).

In this paper, we focus on the challenges specific to the semantic extraction in SC. In Table I, the primary semantic extraction approaches are discussed along with the associated problems that have been solved, partially studied and impeding progress. Some researchers [2], [22], [23], [24], [25], [26] introduced deep learning enabled extraction methods known to be successful in the field of computer vision, natural language processing and speech processing. They have improved robustness in the low SNR regime. However, due to the lack of differentiable loss function, deep learning methods have inherent limitations. In other words, all of these works [2], [22], [23], [24], [25], [26] train neural networks using cross entropy and mean square error, which keeps the present efforts far from the intended practicability over SC [28]. The semantic extraction are handled as a black box in such end-to-end designs, where the semantic channel encoders and decoders must be concurrently trained. Indeed, the efficiency of deep learning enabled extraction is difficult to gauge due to the lack of interpretability.

In different contexts, reinforcement learning has been seen as a viable approach for dealing with the challenges associated with user-defined, task-specific, and perhaps even non-differentiable metrics [27], [28], [29], [30]. As summarized in Table I, learning the appropriate policy through interactions with the surroundings overcomes the limits of semantic metrics’ differentiability [30]. However, it also increases training complexity [28]. Training such a large model from start for a high-dimensional job becomes a significantly challenging task [7]. Furthermore, semantic information will be distinct for different tasks and
therefore deep and reinforcement learning based methods are not applicable for multitask scenarios (may introduce irrelevant information and needs to be retrained, compromising system performance and resources) [31], [32], [33].

Knowledge-based approaches [31], [32] solve multitasking challenge but it is hard to generalize (due to the lack of explainable AI required for extracting semantic information hidden in deep nets). Furthermore, all of them (deep learning based, reinforcement learning based and Knowledge base driven) [2], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33] are impractical for scenarios when semantics are time varying [34], [36], [41].

The semantic-native [34] and probabilistic logic [20] SC system (with relevant insights from [20], [34], [35], [36]), indeed, has a high level of flexibility and adaptability, which is more in line with the evolving conceptualization of 6G. However, the studies [20], [34], [36], once again, are preliminary and theoretical, and putting it into networks remains a tremendous task. With relevant insights from these works (c.f. Table I) understanding their implementation challenges and witnessing the practical success of hash-based signatures such as XMSS [42] in the postquantum cryptography, we take a major departure to exploit the capabilities one time signature (OTS) [43] by developing a hash-based semantic extraction for SC.

b) Effectiveness of Metrics: Research on semantic measurements is relatively poor and most SC assessments are primarily concerned with semantic error, age of information, and its value. However, it is currently unclear how to estimate the quantity of semantic information and the required network resources. See [7], [44] for further details.

c) Quantification of Semantic Noise: Semantic noise is a resultant of the mismatch between the sender and receiver’s knowledge libraries. Most SC studies assume that the sender and receiver share the same invariant knowledge set [2] and the libraries of both parties remain consistent [31]. However, each transceiver’s background knowledge might grow independently, and data can even be intentionally attacked during download and transmission, leading to privacy and security issues. Therefore, relaxation of this fundamental assumption seems mandatory for further research [14].

C. Semantic Extraction: Our Idea and Contributions

Hash functions transform arbitrary-length messages to fixed-length signatures. We mandate cryptographic hash-functions to be one-way (difficult to locate an ingredient in the signature of the input) and collision resistant (hard to discover two inputs that map to the same output). Hashing is one of the most successfully used cryptographic tools, with range of applications. While hashing is employed to handle arbitrary length messages in all practical signature schemes, it has been known (from the beginning of public key cryptography) that they can be used as sole building blocks for secure and meaningful communication. NIST has released SP 800-208 [45] and standardized hash-based signature methods: LMS [46] and SPHINCS+ [47].

Supervised hashing has been known to be successful for document and multimedia retrieval and representation learning [48]. Therefore, we develop a lightweight supervised hashing framework to enable flexible semantic extraction in SC of heterogeneous messages (texts, audio, images, video etc.) by generating a semantics signature.

Recent deep learning-enabled SC studies rely on shared background knowledge between sender and receiver, which is not practicable. We are inspired by the idea proposed in [17] to communicate via discrete tokens derived from a continuous learned space. However, a considerable difficulty of SC lies in handling the discrete constraints imposed on the pursued semantic coding.

As illustrated in Fig. 1, we develop an UBT approach, which consists of i) understand-phase, transforming ‘semantics understanding’ of the message into a signature before transmission by using knowledge base rules to convey the semantic information (sender wants to share with the receiver); and ii) talk-phase, where the generated signature is either transmitted solely or piggybacked with the message, which will be used for reconstructing semantics and adapting knowledge base at the receiver (based on the shared knowledge base, ontology and rules). Unlike the existing works, our learning mechanism is tightly coupled with hashing to generate optimal signatures for transferring semantics from sender to the receiver.

Our main contributions in this paper are outlined as follows:

- We design a simple semantic communication framework based on supervised hashing using signatures. Our novelty in the design of semantic extraction and domain adaptation exemplifies the joint optimization of (information gathering, dissemination, and decision-making) over SC-based 6G networks.
- We develop a novel framework to transform the underlying classification context into an optimization problem and then relax the problem using regularization for generating semantics signatures for SC.

| Semantic extraction methods       | Resolved issues                                                                 | Impeding Challenges                                                                 | Refs.               |
|----------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|---------------------|
| Deep Learning based              | Improved robustness over erroneous links;                                       | Partially studied semantic coding                                                   | (2), (22)–(26)      |
| Reinforcement Learning based     | Developed generalized framework;                                                | But no semantic understanding;                                                      | (27)–(30)          |
| Knowledge Base driven            | Goal-oriented semantic extraction                                               | Poor flexibility in the semantic extraction;                                        | (31)–(33)          |
| Semantic Native and              | Developed contextual reasoning capabilities                                      | Largely theoretical with implementation issues                                      | (20), (34)–(36)    |
| Provenance-based approach        | Proposed flexible, multimodal secure and variable goal solution                 | Implementation focused partially studied details                                    | This work           |
|                                 |                                                                                 |                                                                                      |                     |
New insights developed in this work can be further extended to extensively investigate the benefits for performance, security and privacy of the emerging 6G networks.

II. PRELIMINARIES AND STATE-OF-THE-ART

Shannon’s Information Theory (SIT) states that the transcription between message and symbol usually comprises source encoding that reduces the redundancy contained by the message, followed by channel encoding (structured redundancy to improve reliability). The concepts of SIT govern the rules and properties of source and channel encoding. This combination of source and channel encoding is called syntactic encoding [4], as it changes the form of the message (but not the semantic content of the message). The sequence of symbols then is mapped into a physical electromagnetic wave or an acoustic wave, such that it is suitable for a physical transmission medium.

Traditional source messages in the SIT sense can be applied to encode semantic information, and there is a relationship between syntactic symbols and semantic messages [49]. To elaborate, we consider semantic encoding at a source, which is randomly generating messages with probability \( P_M(m), m \in M \), where \( M \) represents the set of messages. Message \( m \) is produced based on the ontology (and the rules from the knowledge base, see Fig 1), which is converted into a sequence \( x \in X \) of symbols, such that, \( x = f(m) \). The (syntactic) message entropy of the source can then be estimated (by applying Shannon’s theory) as follows:

\[
H(M) = - \sum_{m \in M} P_M(m) \log_2 P_M(m),
\]

(1)

while the logical probability that the source generates the syntactic symbol \( x \) is given by

\[
P(x) = \sum_{x = f(m), m \in M} P_M(m).
\]

(2)

Example 1 Let \( X = \{1, 2, 3\} = \{\text{car, cat, man}\} \). Then, \( M \) is given by a set of images (top right panel of Fig. 1), \( m \) is an image that can be represented by a vector. Therefore, the semantic entropy associated with the symbols generated by the source can be computed as [4]

\[
H(X) = - \sum_{x \in X} P(x) \log_2 P(x).
\]

(3)

Since \( H(X) \) and \( H(M) \) are different, considering the conditional entropies of \( X \) to \( M \) and \( M \) to \( X \), we can efficiently compute the mutual information between \( X \) and \( M \) as

\[
B(M; X) := H(M) - H(M|X) = H(X) - H(X|M),
\]

(4)

where \( H(X|Y) \) is the conditional entropy of \( X \) given \( Y \). As shown in Fig. 1, the semantic coder maintains the meanings linked to the produced syntactic messages. It was partially studied in [50] and reported that an encoder requires on average \( B(M; X) \) bits only, as elaborated in (4), to encode a semantic aspects of a syntactic message generated by the source. For example, the desired block coder for a semantic encoder at the source, as discussed in [50], can be attained simply by employing \( B \) bits semantics per syntactic message without any semantic error, which is referred to as semantic capacity. For simplicity, we omit \( (M; X) \) and use only \( B \). It is worth noting that these \( B \) bits enable us to piggyback semantic signature with the syntactic message packets during transmission.

A. Literature

Of particular importance to this work are the initial research works [2], [13], [14], [15], [16], [17] on SC systems (see the comparison in Table II). Guler et al. [14] initialize research on SC for text information to mitigate semantic errors. Weng et al. [13] developed a deep learning-based system to cope with various channel environments. Xie and Qin [15] applied deep learning to develop distributed SC system for text transmission. Weng and Qin [16] proposed a mechanism to recreate the source message by constructing the speech signal. The authors in [17] developed a neural agent that enables SC via discrete tokens extracted from continuous learning space.

We first formulate an optimization problem of the learning-based hashing by considering the underlying classification problem essential to generate semantics signature (see Section III A). After that, we exploit the regularization concept from large-scale optimization to solve the relaxed optimization with the discrete binary constraints (Section III B), which helps us solve the problem iteratively (Section III C, Algorithm 1).

III. HASHING-BASED SEMANTIC COMMUNICATION

In 6G, we expect to support SC of heterogeneous messages by generating a semantics signature through semantic encoding. For semantic encoding, the task of learning binary hash codes is essential for the proposed framework to find semantics signatures. This task can be carried out by formulating a classification problem. In this section, we first provide the significance of the proposed framework and then formulate a classification problem and derive an iterative approach to solve the problem.

A. Proposed SC Framework

We introduce a new concept, a major departure from the traditional approach to designing and evaluating communication networks, by incorporating information semantics. The information semantics in our design are defined not simply as the

| Table II Comparing Relevant Research Works on Semantic Extraction |
|-----------------|-----------------|-----------------|-----------------|
| Refs | Semantics | Channel | Approach |
| [2] | Text (Sentences) | AWGN | Deep Learning |
| [14] | Text (Words) | binary symmetric | Bayesian game |
| [13] | Speech Signal | AWGN, Rayleigh | Deep Learning |
| [15] | Text | Rayleigh, Rician | Deep Learning |
| [16] | Speech Signal | AWGN, Rayleigh | Deep Learning |
| [17] | Continuous space | Noisy agent | Neural Agent |
| This work | Text, Audio Images, Video | AWGN | Discrete Hashing/Supervised Learning |

POKHREL AND CHOI: UNDERSTAND-BEFORE-TALK (UBT): A SEMANTIC COMMUNICATION APPROACH TO 6G NETWORKS 3547
meaning of the messages, but rather as their significance or importance, within a real-time constraint, and feasibly attributable to the goal of the communication. Our proposed signatures are the keys to i) efficiently synchronize the communication at all three levels and ii) consistently trigger updates of the knowledge bases. The signature employs a set of short binary codes to encode texts, images, videos, and other multimodal data while maintaining their original similarity. By using the signatures, due to the remarkable efficiency of pairwise comparison using the Hamming radius, the task of the nearest neighbour search can be readily performed on a large-scale knowledge base. Furthermore, such learned compact signatures (binary codes generated using data-driven hashing methods) can effectively and highly efficiently index and organize massive data in the knowledge base.

Importantly, the signatures learned can be viewed as non-linearly synthesized feature vectors of the original data, which can jointly optimize information collection, distribution, and decision-making policies for a large scale autonomous network systems. The label information can be used to classify these feature vectors. One can observe in the proposed framework, how our design non-linearly converts the original data into a binary space and then categorize the context-aware similarity in this space. By piggybacking the signatures with the message, we are synchronizing the communication at all three levels and advocating that the joint optimization (information gathering, information dissemination, and decision-making) is essential for the SC. The signature at the receiver side serves as a key to i) connect semantic decoder with the knowledge base, ii) update the knowledge base (if needed) and iii) adjust the hashing function at the decoder to generate the original semantics. Next, we discuss knowledge-base and semantic medium in detail.

Knowledge Base: Facts, relations, and possible ways of reasoning are examples of aspects of real-world knowledge that can be understood, recognised and learnt by all communication parties – forming the core elements of the Knowledge Base. It is worth noting that the users at the source and destination don’t always have to share the same knowledge base. However, both communication parties (sender and receiver) must be aware of the knowledge elements involved in their communication. To capture the rich semantics of knowledge elements as well as their complicated interaction, the sender and receiver must constantly maintain and continuously update their knowledge of the area models. As mentioned earlier, the updates of the knowledge-base in the proposed framework can be mostly triggered by the piggybacked signature while reconstructing semantics at the receiver end.

Semantic Medium: There are primarily two kinds of transmission mediums in the proposed SC system. The first kind of medium is the physical wired/wireless medium, the so-called ‘transmission medium’ where impairments on the transmitted signals are prevalent, due to fading, shadowing, distortion, and interference. Historically, most research efforts in communication have focused on alleviating the adverse effects of such impairments over the physical transmission medium. There exists a second kind of medium in SC, referred to as ‘semantic medium,’ which is very different from the physical transmission medium. It doesn’t focus on transmitting exactly the same message conveyed by the sender but rather aims to preserve the original meaning of the message.

An example of how semantic medium works is shown in Fig. 2. Observe in Fig. 2, although semantic medium does not accurately transmit Shiva’s and Susie’s original messages, it always preserves the sender’s main intention. Therefore, if we evaluate the underlying communication success, there can be a technical engineering failure (at level A), but there is no semantic failure (at level B). Semantic medium can be adversely impacted by the semantic noise generated due to misinterpretation mistakes, or disruption in the estimation, retrieval or prediction of the information, which is one of the challenging problems in the SC. Perhaps the piggybacked signature can be used as a tool for ameliorating semantic noise in the medium as the label information can be used to classify their feature vectors.\footnote{The observation of using signatures as a tool for detection and correction of semantic noise in the medium requires further investigation and is left for future work.}

### B. Learning Signatures by Fast Supervised Learning

In recent years, real momentum has been building up for developing learning-based hashing techniques because they support efficient processing and retrieval for higher-dimensional data such as images, videos, documents, etc. In the proposed SC framework, we often deal with high-dimensional data. Therefore we adopt classification information to learn binary hash codes (i.e., semantics signatures). We consider a discrete hashing mechanism consisting of \( n \) message samples \( \mathbb{X} = \{x_i\}_{i=1}^{n} \), where our objective is to learn a set of binary codes \( \mathbb{B} = \{b_i\}_{i=1}^{n} \) to preserve their similarities in the semantics signature \( b_i \) of \( B \) bits. More importantly, to leverage supervised learning, we inherit binary coding in linear classification with a hypothesis that optimal binary codes are beneficial for such classification problems.

We adopt vectors \( w_k \in \mathbb{R}^{B \times 1} \) to perform the classification for class \( k \in \{1, 2, \ldots, c\} \), and formulate the underlying classification problem as follows:

\[
\min_{\mathbb{B}, \mathbb{Y}, \mathbb{W}} \sum_{i=1}^{n} \text{Loss}(b_i, y_i^T) + \alpha \|W\|^2, \tag{5}
\]

where \( \mathbb{Y} = \{y_i\}_{i=1}^{n} \) is the ground truth matrix (if \( x_i \) is of class \( k \), \( y_{ki} = 1 \) otherwise 0), \( \text{Loss}(b_i, y_i^T) \) is the loss function, \( \alpha \) is
the regularization parameter, and $\mathbb{W} = [w_1 \cdots w_c]$ as
\[
\sum_{i=1}^{n} \text{Loss}(b_i, y, \mathbb{W}) := \|B - \mathbb{W}\|^2.
\] (6)

Observe in (5) that we need
\[
b_i = \text{sgn}(F(x_i)), i \in \{1, 2, \ldots, n\},
\]
even though it appears that we are only interested in learning the hashing function $F(x)$. Moreover, along the lines of [51], removing $b_i$ for eliminating auxiliary variables and constraints is not possible here (as $b_i$ is of central focus). Also, such an approach leads to increasing computational complexity, which becomes not only very difficult but also more sluggish to minimize (5). In particular, $\mathbb{H} = \text{Hash function}$ is the Hash function for encoding $x$ by $B$ bits, where $\text{sgn}(\cdot)$ outputs 1 for all +ve numbers, otherwise $-1$.

With the ideas from [52], we developed a significantly more effective and efficient approach as solving (5) is complicated (NP-hard) even with discrete $b_i$. A simple relaxation to this end would be obtained by assuming that $b_i = F(x_i)$ is continuous, which is applied in most existing hashing schemes. However, the obtained solutions are sub-optimal, although this relaxation simplifies the complexity and makes (5) easier to solve. Therefore, we exploit the different ideas of *regularization from large-scale optimization* using a primal-dual approach to solve the problem more efficiently by keeping the discrete binary constraints of the semantic signature $b_i$ [53].

C. Relaxation Using Regularization

Based on the above discussion and insights from [54], by using penalty parameter, $P$, (5) can be approximated as follows [52]:
\[
\begin{align*}
\min_{B, W, H} & \quad \left( P \sum_{i=1}^{n} \|b_i - F(x_i)\|^2 \\
+ & \sum_{i=1}^{n} \text{Loss}(b_i, y, \mathbb{W}) + \alpha \|\mathbb{W}\|^2 \right) \\
\text{subject to} & \quad b_i \in \{-1, 1\}^B.
\end{align*}
\] (7)

We can see that (7) is non-convex and still highly difficult to minimize. Therefore, for tractability, we adopt a fixed point formulation where we keep two-variable fixed while computing one and thus solve (7) one by one in an iterative fashion. To this end, we develop an iterative procedure, and the design details are provided next (see Algorithm 1).

IV. SIGNATURE LEARNING ALGORITHM AND ANALYSIS

The important design details of Algorithm 1 are explained as follows. By using the fixed-point formulation, for given $B$ in (7), $Q$ is the projection matrix that can be computed by regression and is (Algorithm 1, step 11) given by
\[
Q = (\phi(X)\phi(X)^{\top})^{-1}\phi(X)B^{\top}.
\] (8)

We find that any suitable linear or nonlinear learning algorithm can be used for $F(x)$. For our evaluation, we use a simple nonlinear $F(x) = Q^T\phi(x)$, where $\phi(x)$ is a vector realized using radial basis kernel function of width $d$:
\[
\phi(x) = \begin{bmatrix}
\exp\left(-\frac{\|x - x_1\|^2}{d}\right), \ldots, \exp\left(-\frac{\|x - x_m\|^2}{d}\right)
\end{bmatrix}^{\top}.
\]

Here, $\{x_i\}_{i=1}^{m}$ is obtained from a random selection of $m$ samples from the training set $\{x_i\}_{i=1}^{n}$ (see Algorithm 1, Steps 5–6). In Algorithm 1, step 10, for a given $B$, we compute the classification $\mathbb{W}$ by using multiclass kernel-based support vector machine (SVM) formulation of (7). In steps 12–13 of Algorithm 1, we reconsider a fixed point formulation accounting all variables with hinge loss, but $b_i$ is fixed. After relevant simplification, we observe that (7) can then be expressed as
\[
\begin{align*}
\min_{b_i} & \quad \|b_i - F(x_i)\|^2 - \frac{1}{P} \sum_{k=1}^{c} (w^{(ki)})^{\top} y^{(ki)} + b_i \\
\text{subject to} & \quad b_i \in \{-1, 1\}^B \quad \text{where} \quad w^{(ki)} = w_i - w_k \\
& \quad y^{(ki)} = y_k + s_i - 1, \quad \text{and} \quad s_i \geq 0 \quad \text{is a slack variable}. (9)
\end{align*}
\]

Therefore, revising (8), we get
\[
\begin{align*}
\max_{b_i} & \quad b_i^{\top} \left( F(x_i) + \frac{1}{2P} \sum_{k=1}^{c} w^{(ki)} \right) \\
\text{subject to} & \quad b_i \in \{-1, 1\}^B, \quad (10)
\end{align*}
\]

which has the optimal solution at (Algorithm 1, Step 13)
\[
b_i = \text{sgn} \left( F(x_i) + \frac{1}{2P} \sum_{k=1}^{c} w^{(ki)} \right).
\]

\[\text{We use LIBLINEAR – a library for large linear classification, available at https://web.stanford.edu/~xLIN/liblinear.}\] The hinge loss is considered in loss function for training classifiers, in particular for the SVM.
A. Theoretical Analysis of Algorithm 1

Prior to presenting our theoretical analysis, we revise (7) to guarantee invariant regularization parameters with respect to code length $B$, class size $k$ and $n$ sample instances. Then, the resulting objective function becomes

$$
\min_{\mathbf{b}, \mathbf{W}, \mathbf{t}} \left\{ \frac{P'}{nB} \| \mathbf{b} - F(\mathbf{X}) \|^2 + \frac{1}{nB} \| \mathbf{b} - \mathbf{Y} \mathbf{W} \|^2 + \alpha' \frac{k}{B} \| \mathbf{W} \|^2 \right\},
$$

(11)

Observe that (11) becomes (7), when $P' = P$ and $\alpha' = \alpha k/n$. Now, we compute the computational complexity of Algorithm 1 and then develop its convergence analysis.

Computational Complexity. It is worth noting that our hashing-based SC is a lightweight approach (compared to deep learning frameworks [2], [13], [15], [16]). With our training set consisting of $n$ instances, where each instance is of $d$ dimension feature, and $k$ is the type of categories in the training set. Given $B$ bits code, the time complexity for learning: i) $\mathcal{B}$ is $O(ndB + nkB^2)$; ii) $\mathcal{W}$ is $O(ncB + nB^2 + nk^2 + B^3)$; and iii) $\mathcal{Q}$ is $O(ndB + 2nd^2)$. Therefore, total training time complexity is $O(ndB + nkB^2 + B^3)$.

Using the ideas from [53], [55], we have the following convergence analysis of Algorithm 1.

Theorem 1: In the $t^\text{th}$ iteration, for given $\mathcal{B}_t$ and $\mathcal{F}_{t-1}$, Algorithm 1 converges when learning $\mathcal{W}_t$. After the $t^\text{th}$ iteration, let $\mathcal{W}_t(\mathbf{X}_t)$ and $\mathcal{W}_t(\mathbf{X}_2)$ are learned by using $\mathbf{X}_t$ and $\mathbf{X}_2$ respectively, then assuming $\| \mathbf{b} - \mathbf{y} \mathbf{W} \|^2 \leq \varepsilon$, where $\varepsilon$ is a constant, we have

$$
\| \mathcal{W}_t(\mathbf{X}_t) - \mathcal{W}_t(\mathbf{X}_2) \| \leq \frac{2\varepsilon}{\alpha' n}.
$$

(12)

Proof: We apply the properties of Bergman divergence, denoted by $D_g(\mathbf{X}, \mathbf{Y})$, and its extension [55] to prove Theorem 1 [53]. Given a differentiable convex function that maps matrices to the extended real numbers, the Bergman divergence is always additive and non-negative. Mathematically, $D_g(\mathbf{X}, \mathbf{Y})$ can be computed as [55]

$$
D_g(\mathbf{X}, \mathbf{Y}) := g(\mathbf{X}) - g(\mathbf{Y}) - \langle \nabla g(\mathbf{X}), \mathbf{X} - \mathbf{Y} \rangle,
$$

(13)

where the inner product $\langle \mathbf{X}, \mathbf{Y} \rangle = \Re\{\text{Tr}(\mathbf{X} \mathbf{Y}^\dagger)\}$. In the $t^\text{th}$ iteration, let

$$
\begin{align*}
\mathcal{F}_t(\mathcal{W}) &= \frac{P'}{nB} \| \mathbf{b} - F_{t-1}(\mathbf{X}) \|^2 + \frac{1}{nB} \| \mathbf{b} - \mathbf{Y} \mathbf{W} \|^2 + \alpha' \frac{k}{B} \| \mathbf{W} \|^2, \\
\mathcal{G}_t(\mathcal{W}) &= \alpha' \frac{k}{B} \| \mathbf{W} \|^2.
\end{align*}
$$

Applying Bergman divergence with $\mathcal{F}_t(\mathcal{W})$, after simplification, it can be shown that

$$
\mathcal{D}_{\mathcal{F}_t}(\mathcal{W}_t(\mathbf{X}_2), \mathcal{W}_t(\mathbf{X}_1)) + \mathcal{D}_{\mathcal{F}_t}(\mathcal{W}_t(\mathbf{X}_1), \mathcal{W}_t(\mathbf{X}_2))
= \frac{1}{nB} \left[ \| \mathbf{b}_t - \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_1)^\dagger \|^2 - \| \mathbf{b}_t - \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_1)^\dagger \|^2 + \| \mathbf{b}_t - \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_2)^\dagger \|^2 \right] + \frac{2\varepsilon}{nB} \| \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_1)^\dagger - \mathcal{W}_t(\mathbf{X}_2) \|
\leq \frac{2\varepsilon}{nB} \| \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_1)^\dagger - \mathcal{W}_t(\mathbf{X}_2) \|
+ \frac{2\varepsilon}{nB} \| \mathbf{y}_t \mathcal{W}_t(\mathbf{X}_1)^\dagger - \mathcal{W}_t(\mathbf{X}_2) \|
\leq \frac{4\varepsilon}{nB} \| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \|. 
$$

(14)

Similarly, by applying Bergman divergence with $\mathcal{G}_t(\mathcal{W})$ after simplification

$$
\mathcal{D}_{\mathcal{G}_t}(\mathcal{W}_t(\mathbf{X}_2), \mathcal{W}_t(\mathbf{X}_1)) + \mathcal{D}_{\mathcal{G}_t}(\mathcal{W}_t(\mathbf{X}_1), \mathcal{W}_t(\mathbf{X}_2))
= 2\alpha' \frac{k}{B} \| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \|^2.
$$

(15)

And we know that [55],

$$
\mathcal{D}_{\mathcal{F}_t}(\mathcal{W}_t(\mathbf{X}_2), \mathcal{W}_t(\mathbf{X}_1)) + \mathcal{D}_{\mathcal{F}_t}(\mathcal{W}_t(\mathbf{X}_1), \mathcal{W}_t(\mathbf{X}_2))
\geq \mathcal{D}_{\mathcal{G}_t}(\mathcal{W}_t(\mathbf{X}_2), \mathcal{W}_t(\mathbf{X}_1)) + \mathcal{D}_{\mathcal{G}_t}(\mathcal{W}_t(\mathbf{X}_1), \mathcal{W}_t(\mathbf{X}_2))
\geq \frac{4\varepsilon}{nB} \| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \|^2,
$$

which proves that

$$
\| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \| \leq \frac{2\varepsilon}{\alpha' n}.
$$

(16)

It is clear that Theorem 1 states that $\| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \|$ (and $\| \mathcal{B}_t(\mathbf{X}_1) - \mathcal{B}_t(\mathbf{X}_2) \|$) decreases with $n$. In addition, we can see that in each iteration the proposed loss $\| \mathbf{b} - \mathbf{Y} \mathbf{W} \|^2$ mandates each signature $\mathbf{b}_t$ of $\mathbf{x}_t$ is from $\mathcal{W}$ using $\mathbf{y}_t$. Also our proposed signature learning algorithm is stable with respect to $\mathcal{W}$. Our design of signature learning mandates that the gap $\| \mathcal{W}_t(\mathbf{X}_1) - \mathcal{W}_t(\mathbf{X}_2) \|$ and the difference $\| \mathcal{B}_t(\mathbf{X}_1) - \mathcal{B}_t(\mathbf{X}_2) \|$ decreases when $n$ increases and therefore the convergence of Algorithm 1 is guaranteed.

B. With Domain Adaptive Hashing (DAH)

Using ideas from [56], [57], we extend our hashing-based SC approach for domain adaptation that includes unsupervised domain adaptation between the sender’s and receiver’s knowledge base, semantic extraction at the sender, and unsupervised hashing for the receiver. We employ Multi-Kernel Maximum Mean Discrepancy (MKMMD) [57] to quantify the distribution difference between sender and receiver datasets in a reproducing-kernel Hilbert space which help reducing knowledge base disparity through nonlinear data alignment. In particular, with the extended framework, we seek to minimize (slight abuse of notations to distinguish sender and receiver and extending the ideas from [57])

$$
\min_{\mathbf{x} \in \{\mathbf{X}_s \cup \mathbf{X}_r\}} \mathcal{J} = F(\mathbf{X}_s) + \eta H(\mathbf{X}_s, \mathbf{X}_r) + \gamma \mathcal{D}_A(\mathbf{X}_s, \mathbf{X}_r)
$$

(17)

where $(\eta, \gamma)$ pair decides the importance of the entropy loss, $H(\mathbf{X}_s, \mathbf{X}_r)$ and the domain adaptation, $\mathcal{D}_A(\mathbf{X}_s, \mathbf{X}_r)$. We follow the heuristics explained in [57] to set $(\eta, \gamma)$ in our evaluation.
Note that the hash values in the extended approach are computed from the output of the whole network, and \( \mathbb{H} = \text{sgn}(F(\{x_s, x_r\})) \) is the extended Hash function for encoding \( x \) by \( B \) bits. With \( \mathcal{X}_s = \{x_1, \ldots, x_n\} \) and \( \mathcal{X}_r = \{x_1, \ldots, x_m\} \), mathematically, \( H(\mathcal{X}_s, \mathcal{X}_r) \) and \( DA(\mathcal{X}_s, \mathcal{X}_r) \) can be computed by using (3) and (8) respectively

\[
H(\mathcal{X}_s, \mathcal{X}_r) = -\frac{1}{n_r} \sum_{i=1}^{n_r} \sum_{j=1}^{n_r} P(ij) \log_2 P(ij) \tag{18}
\]

\[
DA(\mathcal{X}_s, \mathcal{X}_r) = \sum ||E[\phi(\mathcal{X}_s)] - E[\phi(\mathcal{X}_r)]||_2^H \tag{19}
\]

where, \( ||E[\phi(\mathcal{X}_s)] - E[\phi(\mathcal{X}_r)]||_2^H \) is the MKMMD associated with the corresponding layer represented by the kernel Hilbert space \( \mathcal{H} \) of the resultant kernel \( \langle \phi(\mathcal{X}_s), \phi(\mathcal{X}_r) \rangle \).

C. Practical Benefits of the Proposed SC Framework

Applications that rely on real-time decision-making, such as smart grids and networked control, need a major rethink in their communication processes. There is an increasing interest in the investigation of an efficacious SC system in 6G that can jointly optimize information collection, distribution, and decision-making policies for a large scale autonomous network systems [5, 6].

While such network systems are not-so-futuristic, the theoretical foundation for SC is poor since “the proper or meaningful piece of information” is neither specified in classical data communication, nor is the process of information generation considered in the communication process. For the last decade, our objective had been to ultra-reliably deliver an entire data stream as quickly as practicable [6]. However, such a classic strategy can indeed be extremely inefficient for 6G systems. Let us analyze the following two beneficial cases for 6G with intelligent transportation systems and industry 4.0 systems.

- **Case 1**: Negotiating for timely consensus among autonomous cars regarding intended moves is critical for preventing accidents in intelligent transportation systems. This is critical in unexpected situations, such as the abrupt presence of pedestrians. The proposed hashing based SC approach, considering the time-dependent value of messages help the underlying network prioritize information flow while fulfilling safety requirements. An efficient and low latency signature learning mechanism (developed in Algorithm 1) enables the transmission of signatures along with message, which can build the foundation towards the desired goal-oriented communication [5, 19].

- **Case 2**: In industry 4.0 network systems, plant statuses are frequently relayed to the remote server for intelligent manufacturing and process estimation. Communication between plants and servers, on the other hand, might experience unforeseen delays, not desirable in 6G [5]. In this scenario, it has already been demonstrated that employing process-aware sparse sampling rather than uniform sampling can enhance performance by orders of magnitude [2]. This observation justifies our rationale behind designing Algorithm 1 for learning signatures (which can perform context-aware sampling) in 6G and piggybacking them with messages for the desired joint optimization across three levels. Such an approach can dynamically maximize 6G networks’ utility, significantly reduce the sampling rate and the corresponding energy consumption on the plant side, which are vital for low-power and energy harvesting sensors over the industry 4.0 systems.

V. PERFORMANCE EVALUATION

We perform experiments and evaluate the proposed hashing-based SC in terms of computational complexity, efficiency, and reconstruction performance at the receiver end. We have first utilized CIFAR-10. CIFAR-10 consists of 60 k plus images categorized into ten classes, where each class consists of six thousand image samples. In our setup, we segregate the datasets into two parts, where five thousand samples are used for testing and the remaining for the training, which are large image data sets tested extensively in the literature. The samples taken from CIFAR-10 data are normalized based on our requirements in data size (e.g., \( B \) in bits). We set \( \alpha = 1 \), maximum number of iterations \( I = 100 \), and \( P = 0.0001 \).

Several metrics have been developed with a view to assessing the code generated by the hashing methods [48, 58]. The precision and mean average precision (MAP) are compared with different lengths in iterations \( I = 100 \), and \( P = 0.0001 \).

Mathematically, given \( \mathcal{N} \) retrieval queries [48, 58]

\[
\text{precision}(r) = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} \left( \frac{\sum_{r} Y_{j,r}^{+}}{\sum_{r} Y_{j,r}} \right) ;
\]

\[
\text{MAP} = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} \frac{1}{\mathcal{N}} \sum_{k=1}^{n} \text{precision}(k) \sum(y_{ki} = 1).
\]

We have observed the results in precision (look up of \( \text{dist}_H = 2 \)) and MAP. To this end, ground truths are determined from the data sets by using specific category information, and if there is no return while calculating precision, it is considered a false case.

Our first channel model is an emulator that adds white Gaussian noise to the signals passing through it. This AWGN model is applied to each element of the whole input signal where the scalar value of SNR is introduced. We also consider Rayleigh and Rician channel for further evaluations. For a given SNR, the precision and MAP are compared with different lengths in \( B \) (in bits), as illustrated in Figs. 3(a) and 3(b).

Figs. 3(a) and 3(b) show that the proposed algorithm achieves the closest match of precision and MAP at the receiver with ground truth at the sender. The purpose of evaluating both the sender and receiver sides is to quantify the impact of semantic...
Fig. 3. Precision and MAP (\( \text{dist}_{ij} = 2 \)) with the models trained under AWGN channel and compared in terms of semantics signature reconstruction at the receiver. In a) and b), SNR is fixed at 10 dB, \( B \) ranges from 16 bits to 256 bits. However, in c) and d), SNR varies from 6 dB to 18 dB, \( B \) is fixed to 96 bits. (a) precision of semantics reconstruction. (b) MAP of semantics reconstruction (c) precision of semantics reconstruction (d) MAP of semantics reconstruction

noise and compare/observe the reconstruction accuracy. In addition, we have compared the results at both the receiver and sender with one of the well-known hashing methods [51] (the code is publicly available). The later comparison reveals the fact that for higher precision \( b_j \) should be taken into account (recall from Section II-A, that solving (5) is NP-hard with discrete \( b_j \)).

Figs. 3(a) and 3(b) provide an essential insight that one of the main benefits of the proposed hashing-based SC approach is its abstraction efficiency in encoding semantics signatures. Perhaps one of the immediate advancements of the proposed approach could be designed towards generating a single semantic signature to represent even a video, a picture or a sound with much the same semantic, which seems to be invaluable for communication even with ultra-low bandwidth. Furthermore, by fixing \( B = 96 \) bits, the precision and MAP are evaluated with varying SNR of the channel (using AWGN channel model), as illustrated in Figs. 3(c) and 3(d). As noted in Fig. 1, we found that the proposed algorithm achieves the closest precision match in reconstructing semantics and the MAP at the receiver end compared with the ground truth under similar channel conditions. Such semantic reconstruction has an intrinsic capability to offer multimodal experiences in 6G networks, facilitating the transfer of skills over the Internet. We anticipate that such a framework may eventually lead to the emergence of the Internet of Senses (IoS), combining visual, auditory, haptic, and other technologies that enable people to experience feelings/senses remotely.

Extensive evaluations with the models trained under various physical channels (AWGN, Rician, Rayleigh) are shown in Fig. 4. Observe in Fig. 4(a) that the precision of the three models is evaluated over AWGN. The model trained with AWGN outperforms others when SNR is high (> 14 dB). In Fig. 4(b) model trained with AWGN appears poor over the Rayleigh channel. Furthermore, Fig. 4(c) demonstrates their comparison over the Rician channel. Overall, we found that the model trained with the Rician channel setup is resilient to different environments. Fig. 4(d) demonstrates the convergence dynamics of our Algorithm 1 under other channel conditions. The dampings are expected because Algorithm 1 strictly updates each bit (iteratively) till convergence with a better set of semantics signatures, i.e., a set of binary codes \( B = \{ b_j \}_{j=1}^n \). In our evaluations and experiments, we noticed that the whole \( B \) bits could be learned iteratively in \( \tau B \) times, which usually converges when iteration times \( \tau = 3 \) to 6.

A. With Multimodal Dataset

To further evaluate the proposed framework, we used the publicly available multimodal dataset, PKU XMediaNet [59], comprising of five different media (text, image, audio, video and 3D models). To verify the hierarchical semantic structure, Peng et al. [59] applied 200 category nodes from WordNet and generated this dataset consisting of 48 kinds of animals and 152 types of artifacts.

Figs. 5(a) and 5(b) show that the proposed algorithm achieves the closest match of the MAP scores for the audio and the video at the receiver when compared with ground truth at the sender. In addition, we have compared the results at both the receiver and sender by generating signatures with [51] and one of the well-known generative adversarial (GAN) hashing methods [60]. Only for the purpose of benchmarking and to provide a fair assessment, we apply deep reinforcement learning (DRL) features with hashing methods in our experiments. In our
preliminary analysis, we have observed that when our algorithm is integrated with DRL (along the lines of that of [61]), it outperforms [60] in both cases. In our numerical experiments, we found that “Algo+DRL” improved robustness in the low SNR regime, however, due to the lack of differentiable loss function mentioned earlier and real-time communication requirements in the anticipated SC for 6G systems, the adoption of deep hashing-based signatures remains questionable.

**Fig. 4.** Precision and convergence of Algorithm 1 ($B$ is 96 bits and $dist_H$ is 2) with the models trained under various channels and evaluated under AWGN, Rician and Rayleigh channel and the Convergence of Algorithm 1 (loss evolution vs iteration times). (a) Precision over AWGN Channel. (b) Precision over Rician Channel. (c) Precision over Rayleigh Channel. (d) Convergence dynamics of Algorithm 1 (loss vs iteration times).

**Fig. 5.** MAP ($dist_H$ is 2) with the models trained under AWGN channel and compared in terms of semantics signature reconstruction at the receiver. In a) and b), SNR is fixed at 10 dB, $B$ ranges from 16 bits to 256 bits. (a) MAP scores of the audio compared with the semantics reconstruction. (b) MAP scores of the video compared with the semantics reconstruction.

**B. Towards Sophisticated Hashing-Based Signatures**

Our primary focus in this work is to quantify the category information by using semantics signature for the regression or classification to attain the proposed SC by design and therefore we assume negligible correlation between data samples. With relevant insights from this work and by adopting enhanced hashing approaches [60], [61], [62], [63] (studied in a different contexts) an emerging important research direction is towards developing improved signature by capturing the pairwise correlations between samples or by jointly considering the correlations along with semantic reconstruction. For example, one may apply deep image hashing with a deep neural network to convert an input image to compact signatures, allowing efficient image retrieval over large-scale knowledge-base with huge datasets. Deep hashing [62] has attracted increasing attention from the academic community as a result of the enormous increase in the current knowledge base and the data sets. However, the SC needs of proliferating real-time systems in different domains, is not supportable by these paradigms. To analyze this, we performed retrieval efficiency comparison of the different existing methods in the literature. Convolutional neural networks (CNNs), such as ResNet, sit at the heart of deep hashing framework, however, there are several limitations for straightforward applications to SC. To provide a fair assessment, we compare deep learning features retrieved by the CNN network with typical hashing methods in our experiments. The
histograms of the average computation time of different methods are shown in Figs. 6(a) and 6(b). Peng et al. [61] explained insights toward improving hash function correlation by combining deep reinforcement learning with a sequential learning technique.

Whilst generation of the results in Fig. 6 we have found that CNNs perform well when categorizing very similar images to the dataset. However, if the images contain some degree of tilt or rotation, CNNs typically struggle to classify the image. Such observation has already been justified in the literature [64]. With CNNs, neurons provide all minute detail recognition in the convolutional layer. These high-level neurons then determine if all traits are there. The image is strided to determine whether or not features are present. During this phase, the CNN may lose the knowledge of the composition and position of the components in the image and sends the information to a neuron, which often cannot categorize them effectively.

Fig. 6(b) shows the comparison of CPU utilization in signature learning and the mean reconstruction time in milliseconds using CIFAR-10 by fixing the signature length to 64 bits. When comparing with other methods [51], [60], [61], in Fig. 6, our algorithm outperforms with approximately 40% CPU utilization while learning signatures and the average retrieval time is below 5 ms. Due to aforementioned limitations, the results in Fig. 6, we find that existing deep hashing are not directly applicable for the SC paradigm (freshness, age-of-information and the extreme low latency requirements of 6G [6], [65]). Using the extended approach developed in Section IV-B (Algorithm 1 with DAH), we revisit all important results under the same setting and network conditions. In Fig. 7, we show the results of Fig. 3(a) and Fig. 5(a) using DAH for comparison. Figs. 7(a) and (b) demonstrate that both the precision and MAP of semantic reconstruction at the receiver improves significantly with the proposed DAH. As anticipated in Section IV-B, these observations in Fig. 7 reflect the importance of accounting the entropy loss, \( H(\mathcal{X}_s, \mathcal{X}_r) \) and the domain adaptation, \( DA(\mathcal{X}_s, \mathcal{X}_r) \) in improving the accuracy of semantic extraction for the proposed SC framework.

VI. CONCLUDING REMARKS

Semantic communication framework design is the major departure that is rapidly evolving to facilitate the real momentum towards 6G networks. It aims to jointly optimize information collection, distribution, and decision-making policies in 6G systems. We proposed and evaluated a hashing-based SC framework, where the learning objective is to generate the optimal binary hash codes as semantics signatures for transforming semantics understanding along with the semantic message. We have evaluated the extended framework with domain adaptation using large data sets and demonstrated its precision and effectiveness in image and multimedia transmission. In the future, further advancement of the proposed framework could harness the substantial advances in sensor and actuator technology, such as smell actuation and high-quality haptics.

REFERENCES

[1] C. E. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, no. 3, pp. 379–423, 1948.
[2] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, “Deep learning enabled semantic communication systems,” IEEE Trans. Signal Process., vol. 69, pp. 2663–2675, 2021.
[3] M. Kountouris and N. Pappas, “Semantics-empowered communication for networked intelligent systems,” IEEE Commun. Mag., vol. 59, no. 6, pp. 96–102, Jun. 2021.
[4] E. C. Strinati and S. Barbarossa, “6G networks: Beyond shannon towards semantic and goal-oriented communications,” Comput. Netw., vol. 190, 2021, Art. no. 107930.
[5] S. R. Pokhrel, “Learning from data streams for automation and orchestration of 6G industrial IoT: Toward a semantic communication framework,” *Neural Comput. Appl.*, vol. 34, pp. 15197–15206, 2022.

[6] E. Uysal et al., “Semantic communications in networked systems: A data significance perspective,” *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, Jul/Aug. 2022, doi: 10.1109/MNET.106.2100636.

[7] W. Yang et al., “Semantic communications for 6G future internet: Fundamentals, applications, and challenges,” Jun. 10, 2022, arXiv:2207.00427.

[8] J. Choi, “Data-aided sensing for Gaussian process regression in IoT systems,” *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7717–7726, May 2021.

[9] D. Sperber and D. Wilson, *Relevance: Commun. and Cogn.*, vol. 142, Cambridge, MA, USA: Harvard Univ. Press, 186.

[10] M. Sana and E. C. Strinati, “Learning semantics: An opportunity for effective 6G communications,” in *Proc. IEEE 19th Annu. Consum. Commun. Netw. Conf.*, 2022, pp. 631–636.

[11] Q. Lan et al., “What is semantic communication? a view on conveying meaning in the era of machine intelligence,” *J. Commun. Inf. Netw.*, vol. 6, no. 4, pp. 336–371, 2022.

[12] P. Zhang et al., “Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks,” *Engineering*, vol. 8, pp. 60–73, 2022.

[13] Z. Weng, Z. Qin, and G. Y. Li, “Semantic communications for speech signals,” in *Proc. IEEE Int. Conf. Communicat.*, 2021 Jun. 14, pp. 1–6.

[14] B. Güler, A. Venselaar, and A. Swami, “The semantic communication game,” *IEEE Trans. Cogn. Comput. Netw.*, vol. 4, no. 4, pp. 787–802, Dec. 2018.

[15] H. Xie and Z. Qin, “A lite distributed semantic communication system for Internet of Things,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 142–153, Jan. 2021.

[16] Z. Weng and Z. Qin, “Semantic communication systems for speech transmission,” *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2434–2444, Aug. 2021.

[17] M. Tucker et al., “Emergent discrete communication in semantic spaces,” *Adv. Neural Inf. Process. Syst.*, Dec. 6 2021, pp. 34:10574–10574:86.

[18] S. R. Pokhrel and J. Choi, “Federated learning with blockchain for autonomous vehicles: Analysis and design challenges,” *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4734–4746, Aug. 2020.

[19] S. R. Pokhrel, “Software defined internet of vehicles for automation and orchestration,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3890–3899, Jan. 2021.

[20] J. Choi, S. W. Loke, and J. Park, “A unified view on semantic information and communication: A probabilistic logic approach,” in *Proc. IEEE Int. Conf. Commun. Workshops*, 2022, pp. 705–710.

[21] J. M. D. Nafría and B. M. Al Hadithi, “Are “the semantic aspects” actually irrelevant to the engineering problem,” *tripleC: Commun. Capitalism Critique. Open Access J. Glob. Sustain. Inf. Soc.*, vol. 7, no. 2, pp. 300–308, 2009.

[22] C.-H. Lee, J.-W. Lin, P.-H. Chen, and Y.-C. Chang, “Deep learning-constructed joint transmission-recognition for Internet of Things,” *IEEE Access*, vol. 7, pp. 76547–76561, 2019.

[23] J. Xu, B. Ai, W. Chen, A. Yang, P. Sun, and M. Rodrigues, “Wireless image transmission using deep source channel coding with attention modules,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 4, pp. 2315–2328, Apr. 2022.

[24] Q. Hu, G. Zhang, Z. Qin, Y. Cai, and G. Yu, “Robust semantic communications against semantic noise,” Feb. 7, 2022, arXiv:2202.03338.

[25] N. Farsad, M. Rao, and A. Goldsmith, “Deep learning for joint source-channel coding of text,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2018, pp. 2326–2330.

[26] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, “Deep source-channel coding for sentence semantic transmission with HARQ,” *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5225–5240, Aug. 2022.

[27] F. Carpi, C. Hager, M. Martalò, R. Raheli, and H. D. Pfister, “Reinforcement learning for channel coding: Learned bit-flipping decoding,” in *Proc. 57th Annu. Allerton Conf. Commun. Control, Comput.*, 2019, pp. 922–929.

[28] K. Lu et al., “Rethinking modern communication from semantic coding to semantic communication,” *IEEE Wireless Commun.*, early access, May 09, 2022, doi: 10.1109/MWCL.013.2100642.

[29] X. Kang, B. Song, J. Guo, Z. Qin, and F. R. Yu, “Task-oriented image transmission for scene classification in unmanned aerial systems,” *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5181–5192, Aug. 2022.

[30] K. Lu, R. Li, X. Chen, Z. Zhao, and H. Zhang, “Reinforcement learning-powered semantic communication via semantic similarity,” Aug. 27, 2021, arXiv:2108.12121.
[59] Y. Peng, J. Qi, and Y. Yuan, “Modality-specific cross-modal similarity measurement with recurrent attention network,” IEEE Trans. Image Process., vol. 27, no. 11, pp. 5585–5599, Nov. 2018.

[60] J. Zhang and Y. Peng, “Multi-pathway generative adversarial hashing for unsupervised cross-modal retrieval,” IEEE Trans. Multimedia, vol. 22, pp. 174–187, 2020.

[61] Y. Peng, J. Zhang, and Z. Ye, “Deep reinforcement learning for image hashing,” IEEE Trans. Multimedia, vol. 22, pp. 2061–2073, 2020.

[62] Q. Li, Z. Sun, R. He, and T. Tan, “Deep supervised discrete hashing,” in Proc. 31st Int. Conf. Neural Inf. Process. Syst., Dec. 4 2017, pp. 2479–2488.

[63] Z. Zhang, X. Zhu, G. Lu, and Y. Zhang, “Probability ordinal-preserving semantic hashing for large-scale image retrieval,” ACM Trans. Knowl. Discov. From Data, vol. 15, no. 3, pp. 1–22, 2021.

[64] S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic routing between capsules,” in Proc. 31st Int. Conf. Neural Inf. Process. Syst., Dec. 4, 2017, pp. 3859–3869.

[65] S. R. Pokhrel, L. Pan, N. Kumar, R. Doss, and H. L. Vu, “Multipath TCP meets transfer learning: A novel edge-based learning for industrial IoT,” IEEE Internet Things J., vol. 8, no. 13, pp. 10299–10307, Jul. 2021.

Shiva Raj Pokhrel (Senior Member, IEEE) was born in Damak, Nepal. He received the B.E. and M.E. degrees from Pokhara University, Nepal, in 2007 and 2013, respectively, and the Ph.D. degree in 2017 from Swinburne University of Technology, Melbourne, VIC, Australia. He is currently a Lecturer of mobile computing with Deakin University, Geelong, VIC, Australia. He was a Research Fellow with the University of Melbourne, Melbourne, VIC, Australia, and a Network Engineer with Nepal Telecom (2007–2014). His research interests include multi-connectivity, federated learning, industry 4.0 automation, blockchain modelling, optimization, recommender systems, 6G, cloud computing, dynamics control, Internet of Things and cyber-physical systems and also their applications in smart manufacturing, autonomous vehicles and cities. He was the recipient of the prestigious Marie Curie Fellowship in 2017–2020 and the finalist of the IEEE Future Networks Connecting The Unconnected Challenge (CTUC) in 2021.

Jinho Choi (Fellow, IEEE) was born in Seoul, South Korea. He received the B.E. (magna cum laude) degree in electronics engineering from Sogang University, Seoul, South Korea, in 1989, and the M.S.E. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 1991 and 1994, respectively. He is currently a Professor with the School of Information Technology, Burwood, Deakin University, Melbourne, VIC, Australia. Prior to joining Deakin in 2018, he was a Professor and the Chair of wireless, with Swansea University, Swansea, U.K., and a Professor with the Gwangju Institute of Science and Technology, Gwangju, South Korea. He authored two books published by Cambridge University Press in 2006 and 2010. His research interests include the Internet of Things (IoT), wireless communications, and statistical signal processing. Dr. Choi was the recipient of number of best paper awards including the 1999 Best Paper Award for Signal Processing from EURASIP. He is on the list of World’s Top 2% Scientists by Stanford University in 2020 and 2021. He is currently the Editor of IEEE WIRELESS COMMUNICATIONS LETTERS and a Division Editor of Journal of Communications and Networks (JCN). He was an Associate Editor or Editor of other journals including IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, Journal of Communications and Networks, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and ETRI Journal.