Semantic Index Assignment

Basak Guler Aylin Yener
Electrical Engineering Department
The Pennsylvania State University, University Park, PA 16802
basak@psu.edu *yener@ee.psu.edu*

***Abstract*—Conventional performance criteria for communication networks do not take into account the semantics of the data to be communicated. For example, (word) error rates treat errors between semantically similar words (car and automobile) and semantically distant words (car and computer) equally. In reality, the *meaning* of the message is distorted much less when *automobile* is recovered instead of *computer* when the intended message is *car*. In order to correctly address the performance of a semantic system, a new performance criterion is necessary that takes into account the semantic similarities between recovered words. We study in this paper the index assignment problem with a source that produces semantic messages to develop a better understanding of how their meanings affect the semantic error performance in a noisy communication network, and in particular for networks with queries. To this end, we utilize the semantic distances based on lexical taxonomies as a distortion measure in a communication system. Our findings indicate the need for development of semantics-aware physical systems that allow for better integration of human factors and intelligence within complex systems design.**

I. Introduction

Semantic interactions between sources, humans, computers or web resources, play an important role in the design of intelligent systems. To this end, the question of how to interpret the semantic contents in order to improve the communication performance is one that is worth raising.

A source in a communication system generates as many codewords as the number of distinct symbols in the input alphabet. In the absence of channel noise, the ordering of codewords is irrelevant to error performance. However, in a noisy channel, a binary codeword has different chances of being recovered as each one of the other codewords. Index assignment provides a structured mapping between source symbols and codewords to improve the error resilience of a communication link.

An early example of index assignment is Gray mapping [1] in which two successive codewords differ by one bit. Index assignment has been extensively studied in the field of vector quantization and channel optimized source quantizers [2]–[4] to determine the optimal mapping from a continuous source to a binary vector quantizer for a noisy channel. We introduce in this paper a semantic distortion based communication network and develop an index assignment scheme to minimize the average semantic distortion in presence of channel noise. We utilize the semantic similarities defined in a lexical database [5]–[7] to quantify the semantic error probabilities, and design a codebook in which codewords that are likely to be confused by the receiver to represent semantically close concepts. Changing the order of the codewords does not effect the physical characteristics of the network such as bit error rates. However, it may have significant impact on the semantic errors that occur between the *meanings* of the intended and recovered messages. By judicious selection of the codewords, words with higher semantic similarity values in between can be mapped to codewords with short Hamming distances. Therefore, when the wrong index is received due to channel noise, the recovered word will likely be semantically similar to the intended word. We focus on similarity measures defined over WordNet taxonomies [8].

We illustrate this idea with the following example. A source wants to convey a semantic message, such as a reply to an inquiry, to a destination using a finite language. Each word in the language is indexed using a mapping known by both parties. The source transmits the binary representation of the index corresponding to the intended word. Each bit has a fixed cross-over, i.e., bit error probability due to transmission through the noisy channel. Suppose the source wants to transmit the message *"A car is approaching."* but due to the channel noise it may be recovered by the destination as *"A person is approaching."* or *"A vehicle is approaching."* Our aim is to design the binary codewords such that, in case of a channel error, each word has a higher probability of being recovered as a semantically similar word such as *car* being recovered as *vehicle* instead of *person*. Although in both cases the conventional bit error probability is the same, the recovered words in the second case have closer meanings, which we refer to as a smaller semantic distance.

We show that structured assignment of codeword indices improves semantic error performance in noisy channels. Our scheme is independent of the code structure for any fixed rate coding and can also be combined with forward error correction. Our approach may prove useful in understanding the relations between physical layer communication and semantic aspects of messages that appear in emerging composite, e.g., social and tactical networks. The remainder of the paper is organized as follows: In Section II, we present the similarity measures. Section III introduces the system model and semantic index assignment, followed by the numerical results in Section IV. The paper is concluded in Section V.

II. Semantic Similarity

Semantic similarity measures how similar two *concepts* are. A concept refers to the meaning of a word or a group of

This research is sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-2-0053.

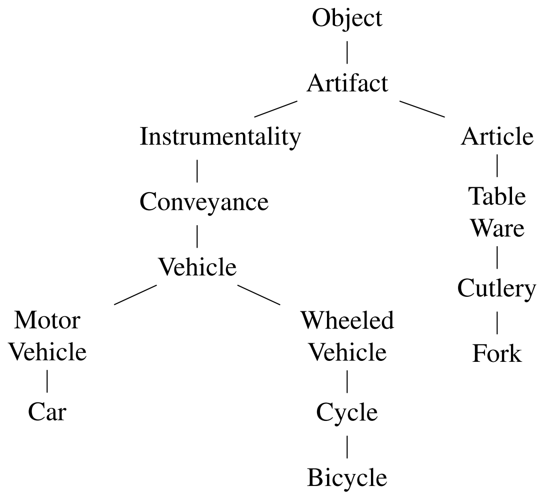

Fig. 1. WordNet taxonomy fragment, [7].

words that exist in a dictionary (corpus) and has a specific meaning in a context. Semantic relations between concepts are often presented by using taxonomies. In that sense, the same word may appear in several places in the taxonomy as the instances of different concepts. Several semantic similarity/distance measures have been formulated in the literature and compared with human experimental results [5].

Word similarity quantifies the semantic similarity between a pair of *words*, with synonyms having the highest value. It is widely used for applications of artificial intelligence, natural language processing and information retrieval. Many well known word similarity measures are based on a thesaurus such as WordNet [8] or statistics from a large corpus [5]–[7], [9] and most semantic applications rely on taxonomy-based measures. A fragment of a WordNet taxonomy used in [5] is presented in Fig. 1. Two main approaches exist for quantifying semantic similarity: node-based and edge-based measures. The node-based similarity measure is introduced in [5] as follows:

$$sim(w_i, w_j) = \max_{c_i, c_j}[sim(c_i, c_j)] \tag{1}$$

where c_i and c_j range over the set of concepts in the taxonomy that are senses of words w_i and w_j, respectively. A single word may have several senses representing different concepts, and may appear in different places in the taxonomy. The notion $sim(c_i, c_j)$ defines the similarity between two concepts:

$$sim(c_i, c_j) = \max_{c \in S(c_i, c_j)}[-log(p(c))] \tag{2}$$

where $S(c_i, c_j)$ is the set of concepts subsuming both c_i and c_j. This measure is based on the information content of lowest common subsumer. A lowest common subsumer in a taxonomy is the concept with shortest distance from two given concepts. For example, animal and mammal are both subsumers of cat and dog, but mammal is lower subsumer than animal. The node-based measure proposed in [6] considers the information content of the lowest common subsumer and the two concepts simultaneously. The edge-based approach [9] utilizes the distance between two nodes in the taxonomy to evaluate the semantic similarity. The hybrid method in [7] takes into account various aspects of the nodes in the taxonomy tree such as information content, depth, degree distribution, and path length.

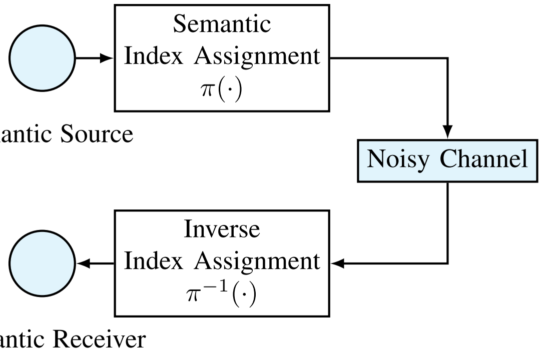
Semantic Index Assignment $\pi(\cdot)$
Semantic Source
Noisy Channel
Inverse Index Assignment $\pi^{-1}(\cdot)$
Semantic Receiver

Fig. 2. Semantic Communication Model.

III. System Model

The semantic communication model we consider consists of a semantic source-destination pair as in Fig. 2. The semantic source chooses words from a finite language, each word defining a distinct source symbol. We consider a set $\mathcal{W}$ of words $\mathcal{W} = \{w_1, \dots, w_{|\mathcal{W}|}\}$. Each source symbol is assigned an index number $i \in \mathcal{I}$ where $\mathcal{I} = \{1, 2, \dots, |\mathcal{I}|\}$ is the set of all indices. The source transmits the binary representation of the index of the intended word through a noisy channel. We define the mapping between the binary representation of the indices and the source symbols by the operation $\pi(\cdot) : \mathcal{W} \to \{0, 1\}^n$ where $n = \log\left(|\mathcal{I}|\right)$, assuming the size of the index set is an integer power of two as often used in practical models. We assume a binary symmetric channel (BSC) with fixed cross-over probability of β. Upon completion of transmission of the binary bit stream, the destination decodes the received binary bit stream to recover the correct word. This is done by an inverse index assignment $\pi^{-1}(\cdot)$. We address in this paper finding the optimal mapping between $|\mathcal{I}|$ indices and $|\mathcal{W}|$ words to minimize the average semantic distance between recovered words. We note that semantic index assignment (SIA) is not necessarily a bijective operation between words and binary codewords, as we show in Section III-B when we discuss index assignment for networks with queries.

A. Semantic Index Assignment

In this section, we assume that each word is assigned to a distinct index, or equally to its binary representation, $|\mathcal{I}| = |\mathcal{W}|$. Therefore, the number of bits required for the binary representation of indices is $\log(|\mathcal{I}|) = \log(|\mathcal{W}|) = n$. The source conveys an intended word $w_i \in \mathcal{W}$ by transmitting the assigned binary codeword $\pi(w_i)$ to the destination through a noisy channel. For a given index assignment $\pi(\cdot)$, the average semantic distortion of the network can be expressed as follows:

$$D(\pi) = \sum_{i=1}^{|\mathcal{W}|} p(w_i) \sum_{j=1}^{|\mathcal{W}|} p(\pi(w_j)|\pi(w_i))d(w_i, w_j) \tag{3}$$

where $|\mathcal{W}|$ is the cardinality of $\mathcal{W}$, $p(w_i)$ is the probability of word w_i, and $d(w_i, w_j)$ is the *semantic distortion* between words w_i and w_j. The probability of receiving the wrong codeword $\pi(w_j)$ for a given $\pi(w_i)$ due to channel errors is:

$$p(\pi(w_j)|\pi(w_i)) = (1-\beta)^{n-h(\pi(w_j),\pi(w_i))}\beta^{h(\pi(w_j),\pi(w_i))} \tag{4}$$

where $h(\pi(w_j),\pi(w_i))$ denotes the Hamming distance between the binary codewords assigned to w_j and w_i. We utilize the normalized semantic similarity measures [6], [7] for evaluating the semantic distortion between any two words:

$$d(w_i,w_j) = 1 - sim(w_i,w_j) \tag{5}$$

where $sim(w_i,w_j) \in [0,1]$ is the normalized semantic similarity between two words w_i and w_j. When comparing two concepts, (5) is modified by $sim(c_i,c_j)$.

The problem we address is to minimize the average total semantic distortion via index assignment:

$$\pi_{opt} = \arg\min_{\pi} \quad D(\pi) \tag{6}$$

in which the distance measure defined in (5) serves to model the semantic distortion between any pair of words. Semantic index assignment (6) is NP-complete, due to the fact that index assignment is a quadratic assignment problem (QAP), which is NP-complete [10]. The QAP nature of the problem in (6) is preserved while using semantic distances.

B. Networks with Semantic Queries

In this section, we investigate the optimal indexing scheme for networks with queries, when the sender is replying to a query made by the other party by using the semantic message set introduced in the previous section. We assume that each query has a limited number of meaningful answers which is a subset of all possible replies. As an immediate example, consider a finite set of words $\mathcal{W} = \{car, room, vehicle, automobile, cat, table, school\}$. The first person makes the inquiry *"What is the moving object approaching the building?"* while knowing that the meaningful answers for the given query are $\{car, cat, vehicle, automobile\}$, and that the words $\{table, room, school\}$ are irrelevant and hence can be ignored. We assume that the inquiries are chosen from a predetermined set $\mathcal{Q}$ known to both parties in advance. Our aim is to utilize the semantic relationships between the words and the queries for index assignment by using the minimum number of codewords. To this end, we start by constructing a characteristic graph [11] that captures the word-query relations. Let $G = (\mathcal{W}, E)$ be a graph on $|\mathcal{W}|$ vertices. We define an edge $(w_i,w_j) \in E$ if the following conditions are satisfied simultaneously for a given query $q \in \mathcal{Q}$:

$$i) \quad P(w_i,q), P(w_j,q) > 0. \tag{7}$$

$$ii) \quad w_i \text{ and } w_j \text{ are not synonyms of each other.} \tag{8}$$

where $P(w_i,q)$ represents the membership condition whether the word w_i is a meaningful reply to inquiry q. That is, $P(w_i,q) > 0$ if w_i is a useful reply to inquiry q whereas $P(w_i,q) = 0$ if it is not. Accordingly, each edge represents a pair of words that correspond to *distinct* meaningful answers for *some* query and therefore need to be distinguished.

Consider the worst-case zero-error transmission of the set of words $\mathcal{W}$ when the word-query relations are given by the characteristic graph $G = (\mathcal{W}, E)$. It follows from [11] that the minimum number of codewords required is the chromatic number χ_G of G. Denote the color $c(w_i)$ of word w_i by the mapping $c : \mathcal{W} \to \mathcal{C}$ where $\mathcal{C}$ is the set of colors $\mathcal{C} = \{1, \ldots, \chi_G\}$. Similarly, binary representations of color indices are mapped to the words through the function $\pi : \mathcal{W} \to \{0,1\}^n$ where $n = \lceil \log(\chi_G) \rceil$ and $\pi(w_i) \neq \pi(w_j)$, $\forall (w_i,w_j) \in E$. That is, each word is given a codeword equal to the binary representation of the color assigned to it.

This problem has two key aspects: achieving a valid χ_G coloring, i.e., partitioning the source space, and achieving minimum semantic distortion by codeword assignment. We note that this problem differs from the index assignment problem considered in the previous section as the assigned binary codewords are no longer distinct. In effect, the solution to this problem requires joint graph coloring and index assignment, which we formulate as a multi-objective optimization problem in the sequel. In order to determine a valid χ_G coloring of the characteristic graph, we utilize the following function:

$$f(c) = \sum_{\substack{(w_i,w_j)\in E \\ w_i,w_j \in \mathcal{W}}} I_{\{c(w_i)=c(w_j)\}}(w_i,w_j) \tag{9}$$

and $I_{\{c(w_i)=c(w_j)\}} : \mathcal{W} \times \mathcal{W} \to \{0,1\}$ is an indicator function:

$$I_{\{c(w_i)=c(w_j)\}}(w_i,w_j) = \begin{cases} 1 & \text{if} \quad c(w_i) = c(w_j) \\ 0 & \text{o.w.} \end{cases} \tag{10}$$

That is, each edge with the same color at both ends incur a cost of 1. It follows that c is a valid coloring if and only if $f(c) = 0$. The above problem is equal to assigning $\lceil \log(\chi_G) \rceil$ binary codewords to $\mathcal{W}$ words while preserving the edge relationships specified by the conditions in (7) and (8) such that no adjacent vertices receive the same codeword. Hence, the first objective function can be formulated as follows:

$$f(\pi) = \sum_{\substack{(w_i,w_j)\in E \\ w_i,w_j \in \mathcal{W}}} I_{\{\pi(w_i)=\pi(w_j)\}}(w_i,w_j) \tag{11}$$

where $I_{\{\pi(w_i)=\pi(w_j)\}} : \mathcal{W} \times \mathcal{W} \to \{0,1\}^{\lceil \log(\chi_G) \rceil}$ is given as:

$$I_{\{\pi(w_i)=\pi(w_j)\}}(w_i,w_j) = \begin{cases} 1 & \text{if} \quad \pi(w_i) = \pi(w_j) \\ 0 & \text{o.w.} \end{cases} \tag{12}$$

The second objective function is defined to find an index assignment to minimize the average semantic distortion caused by the channel for a given coloring. We define the expected semantic distortion of a given index assignment π as follows:

$$D(\pi) = \sum_{\substack{(w_i,w_j)\in E \\ w_i,w_j \in \mathcal{W}}} p(w_i) p(\pi(w_j)|\pi(w_i)) d(w_i,w_j) \tag{13}$$

where $\pi(w_i)$ is the binary representation for the index of w_i:

$$p(\pi(w_j)|\pi(w_i)) = (1-\beta)^{\lceil \log(\chi_G) \rceil - h(\pi(w_j),\pi(w_i))}\beta^{h(\pi(w_j),\pi(w_i))} \tag{14}$$

We note that the receiver can distinguish any two words that are assigned the same codeword by merely using its own inquiry and the conditions in (7)-(8). That is, for each color class,

Algorithm 1 Semantic Index Assignment

```
Choose an initial state for the index assignment $\pi$ randomly.
Define the melting temperature $T_m$ and the freezing
temperature $T_f$.
Initialize the initial temperature $T = T_m$.
Calculate the semantic distortion $D(\pi)$ for state $\pi$.
until $T < T_f$ or a stable state is reached do
    Choose another state $\pi'$ as a perturbation of state $\pi$
    by interchanging two randomly chosen components.
    Calculate the average semantic distortion $D(\pi')$ for
    state $\pi'$.
    Let $\triangle D = D(\pi') - D(\pi)$.
    if $\triangle D < 0$ then
        Set $\pi = \pi'$.
    else
        Set $\pi = \pi'$ with probability $e^{-\triangle D/T}$.
        Lower the temperature.
    end
end
```

Algorithm 2 Joint Graph Coloring and Index Assignment

```
Construct the characteristic graph $G = (\mathcal{W}, E)$.
Choose an initial coloring, though not necessarily a valid
one, by assigning each vertex (word) a codeword from the
set $\{0,1\}^n$ where $n = \lceil log(\chi_G) \rceil$, uniformly at random.
Define melting and freezing temperatures, $T_m$ and $T_f$.
Set the initial temperature to $T = T_m$.
Calculate the cost function $f(\pi)$ using (11).
Determine the semantic distortion $D(\pi)$ from (13).
Let $\phi(\pi) = \alpha_1 f(\pi) + \alpha_2 D(\pi)$.
until $T < T_f$ or a stable state is reached do
    Choose another state $\pi'$ by randomly assigning a
    new binary codeword to a random vertex.
    Let $\triangle\phi = \phi(\pi) - \phi(\pi')$
             $= \alpha_1(f(\pi) - f(\pi')) + \alpha_2(D(\pi) - D(\pi'))$
    if $\triangle\phi < 0$ then
        Set $\pi = \pi'$.
    else
        Set $\pi = \pi'$ with probability $e^{-\triangle\phi/T}$.
        Reduce the temperature.
    end
end
```

there is at most one meaningful answer for each query given that the answers are not synonyms of each other, in which case the receiver can infer the same meaning from both words with no semantic errors. Thus, semantic distortion between words belonging to the same color class, hence assigned the same binary codeword need not be considered in the average semantic distortion calculation. Semantic index assignment for networks with queries is a constrained optimization problem:

$$\begin{aligned} \pi_{opt} = \arg\min_{\pi} \quad & D(\pi) \\ \text{s.t.} \quad & f(\pi) = 0 \end{aligned} \tag{15}$$

Semantic index assignment for networks with queries is NP-complete. Due to space concerns, we only outline that both graph coloring and QAP, which are both known to be NP-complete, can be reduced to index assignment with queries. As (15) is NP-complete, exact search methods become impractical as the word space becomes larger. To overcome this problem, we define a weighted multi-objective optimization problem to find the index assignment for the minimum semantic distortion χ_G coloring of the characteristic graph:

$$\pi_{opt} = \arg\min_{\pi} \quad \alpha_1 f(\pi) + \alpha_2 D(\pi) \tag{16}$$

where α_1 and α_2 are the weights assigned to each function to control its relative importance. In the following section we propose a simulated annealing technique for tackling this multi-criterion optimization problem in practical scenarios.

C. Simulated Annealing for Semantic Index Assignment

The computational complexity of a QAP grows significantly when the set of words grows large, as observed in many semantic applications. Hence, effective approximate methods are proposed in the literature to reduce the complexity in a large search space. We utilize the simulated annealing method [12] which has proved useful for the study of combinatorial problems in addition to index assignment [2], [3], [13].

Simulated annealing is a technique that mimics the physical process of annealing which involves heating a material to its melting point and then slowly cooling it to form crystals at a minimum energy state. The algorithm defines a melting temperature set to a high value at the beginning of the process to provide a high degree of randomness allowing almost all perturbations. As the temperature is decreased, new perturbations are accepted with diminishingly small probabilities, which continues until a sufficiently small temperature, called the freezing temperature, is achieved.

Steps of simulated annealing for the problem considered in Section III-A is provided in Algorithm 1, where each word is assigned a distinct binary bit stream. The new state at each step is achieved by perturbing the current state, and all the states are permutations of one another. However, Section III-B allows multiple words to be mapped to the same binary codeword, and thus the final state, i.e., final assignment, is not necessarily a permutation of the initial state. Thus, a different approach is required in Algorithm 2 for updating the assignments at each step, for which a random vertex of the characteristic graph is assigned a random binary codeword to allow greater flexibility for the states traversed. Simulated annealing has been shown to converge in probability to the global minimum on the condition that the melting temperature T_m is sufficiently large and the cooling schedule is sufficiently slow [14].

IV. Numerical Analysis

This section provides the performance evaluations of the semantic index assignment technique proposed in Sections III-A, III-B, and III-C. Initially, we consider the following set of $|\mathcal{W}| = 16$ words chosen from the set of words studied in the semantic similarity literature [5]–[7] $\mathcal{W} = \{w_1, w_2, \ldots, w_{16}\}$ = {*car, monk, noon, automobile, wizard, lad, midday, jewel, magician, fruit, journey, voyage, food, brother, gem, boy*}. We here note that although the original set contains 28 words, we used 16 words to ensure a valid binary number for $n = \log(|\mathcal{W}|)$. Semantic similarities

TABLE I
PARAMETERS USED IN ALGORITHMS 1 AND 2.

T_m	10
T_f	2.5×10^{-4}
β	0.001
α_1	0.7
α_2	0.3
$\lvert\mathcal{W}\rvert$	8,16
Max. iteration	50000

TABLE II
IMPLEMENTATION OF ALGORITHM 1.

Words	Initial Assignment	Final Assignment
$w_1 \leftrightarrow$ car	1101	0100
$w_2 \leftrightarrow$ monk	0111	1010
$w_3 \leftrightarrow$ noon	1100	1001
$w_4 \leftrightarrow$ automobile	1111	0000
$w_5 \leftrightarrow$ wizard	1001	1111
$w_6 \leftrightarrow$ lad	0011	0111
$w_7 \leftrightarrow$ midday	0110	1101
$w_8 \leftrightarrow$ jewel	1011	0101
$w_9 \leftrightarrow$ magician	1010	1011
$w_{10} \leftrightarrow$ fruit	1110	0010
$w_{11} \leftrightarrow$ journey	0000	1000
$w_{12} \leftrightarrow$ voyage	1000	1100
$w_{13} \leftrightarrow$ food	0010	0011
$w_{14} \leftrightarrow$ brother	0101	1110
$w_{15} \leftrightarrow$ gem	0100	0001
$w_{16} \leftrightarrow$ boy	0001	0110

between any two words are evaluated by the JWS (Java WordNet::Similarity) tool [15]. We use the hybrid semantic similarity measure from [7] with the BNC (British National Corpus). Semantic distortions are calculated from (5).

Algorithm 1 is implemented for the given word set and semantic similarities by using the parameters in Table 1. Since $|\mathcal{W}| = 16$, we use 4 bits to represent each codeword, assigned to a unique word. The initial and final assignments of the codewords from the implementation of Algorithm 1 are presented in Table II. We observe from the final assignment that semantically similar words, as suggested by the human experiment results in [5], such as $\{$*automobile, car*$\}$, $\{$*jewel, gem*$\}$, $\{$*journey, voyage*$\}$, $\{$*midday, noon*$\}$, $\{$*boy, lad*$\}$, are assigned to closer codewords in terms of their relative Hamming distances. We evaluate the performance of simulated annealing in terms of the minimum semantic distortion achieved with respect to the channel crossover probabilities in Fig. 3. In particular, we compare the simulated annealing algorithm with the lower bounds obtained by SDP relaxations [4]. For comparison purposes, we also provide performance results for the following set of smaller dimension $\mathcal{W} = \{$*car, monk, automobile, lad, jewel, brother, jam, boy*$\}$, Assignments leading to the best and worst-case performance can be computed by exhaustive search on this set as presented in Fig 3, in which the optimal values and the upper bound are given in terms of the minimum and maximum average semantic distortion. We consider the joint coloring and index assignment problem from Section III-B. Throughout the analysis, the following set of words are assigned to the vertices of an undirected graph with the given order: $\mathcal{W} = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\} = \{$*car, person, office, automobile, woman, building, vehicle, school*$\}$. We assume that the queries are chosen from the set $\mathcal{Q} = \{q_1, q_2\}$ with the first query given as $q_1 \leftrightarrow$ *What is the moving object approaching the building?* for which the meaningful set of answers are given by the set $\mathcal{A}_1 = \{$*car, person, vehicle, woman, automobile*$\}$, though we note that the queries and the meaningful set of answers are to be defined according to the network interests, which may have different structures in different network models. For the second query, we choose $q_2 \leftrightarrow$ *Where is Bob working at?* with the set of meaningful answers given as $\mathcal{A}_2 = \{$*office, school, building*$\}$. We then construct a characteristic graph $G = (V, E)$ in Fig. 4 with the vertex set $V = \mathcal{W}$ and by using the edge relations from (7) and (8). Note that no edge exits between the synonyms *car* and *automobile*.

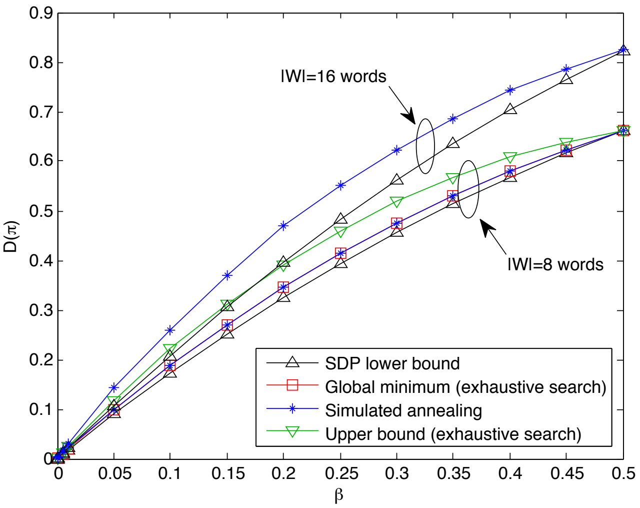

Fig. 3. Average semantic distortion vs. channel crossover probability.

It can be observed from Fig. 4 that the chromatic number χ_G of G is 4. Consider the subgraph induced by the vertices $\{$*car, person, woman, vehicle*$\}$. This subgraph is a complete graph, therefore at least 4 colors are required for a valid coloring. Observe that *automobile* can be assigned the same color as *car* without violating other edge conditions, as the set $\{$*automobile, car, person, woman, vehicle*$\}$ forms the largest connected component of G. The minimum number of bits required for each binary codeword follows from the chromatic number as $\log(\chi_G) = 2$.

We then implement Algorithm 2 using the characteristic graph G with the initial codeword assignments given in Fig. 1. The parameters used in the algorithm are provided in Table I. We use the cooling schedule from [2]. The final codeword assignments are presented in Fig. 6. A closer look at the final codewords gives the following relationships between the subsets of $\mathcal{W}$ and the codewords: 00 $\Leftrightarrow$ $\{$*person, office*$\}$, 01 $\Leftrightarrow$ $\{$*car, automobile*$\}$, 10 $\Leftrightarrow$ $\{$*school, woman*$\}$, 11 $\Leftrightarrow$ $\{$*vehicle, building*$\}$. Note that semantically similar words *car* and *vehicle* are assigned to codewords 01 and 11, with a single bit difference between them. Similarly, *person* and *woman* are assigned close codewords 00 and 10. Synonyms *car* and *automobile* are assigned the same codeword. The algorithm traverses over various valid colorings given in Table III. Although being valid, none of these colorings

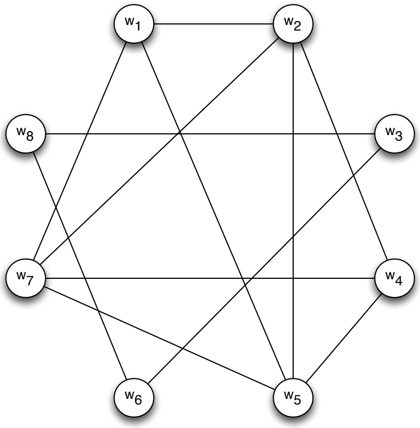

Fig. 4. Characteristic graph G=(V,E).

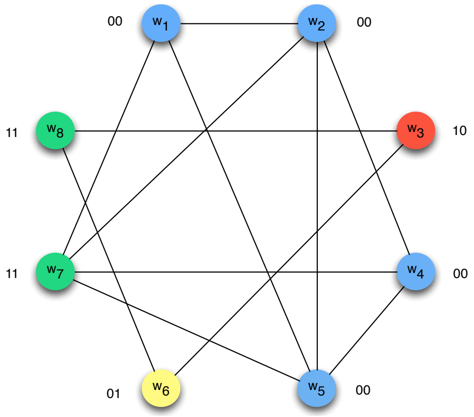

Fig. 5. Initial index assignment for the network with queries.

are final as the algorithm seeks the coloring with the best distortion performance. This is the key factor that makes index assignment for networks with queries challenging, as it cannot be tackled by applying graph coloring and index assignment stages separately. The actual number of valid colorings passed through depend on the system parameters such as the cooling schedule and the freezing temperature. Judicious design of codewords helps to reduce the semantic distortion in noisy and unreliable channel conditions.

V. Conclusion

We have considered the index assignment problem in a query network with semantic sources. We have utilized semantic similarities defined over lexical taxonomies to formulate the semantic distortion between any two words. We have constructed the semantic index assignment problem to achieve minimum average semantic distortion in a noisy communication environment. Future direction is to achieve a unified semantic transmission model with multiple interacting sources.

References

[1] F. Gray, "Pulse code modulation," *United States Patent Number 2633058*, Mar. 1953.

[2] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. on Information Theory*, vol. 36, no. 4, pp. 799–809, Jul. 1990.

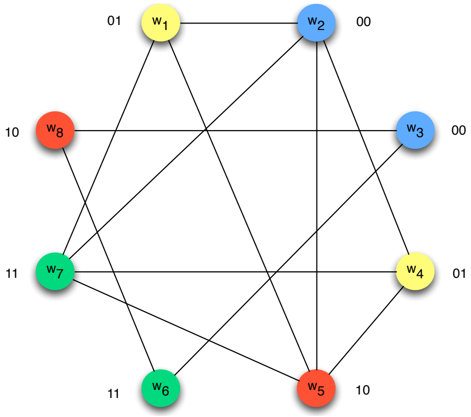

Fig. 6. Final index assignment for the network with queries.

TABLE III
Valid χ_G colorings traversed by Algorithm 2.

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
01	11	10	01	10	01	00	11
01	11	10	01	10	00	00	11
01	11	10	01	10	00	00	01
01	10	11	01	11	10	00	01
01	10	11	01	11	00	00	01
01	10	10	01	11	00	00	01
01	00	10	01	10	00	11	01
01	00	10	01	10	00	11	11
01	00	10	01	10	01	11	11
01	00	00	01	10	01	11	11
01	00	00	01	10	01	11	10
01	00	11	01	10	01	11	10
01	00	00	01	10	11	11	10

[3] S. B. Z. Azami, P. Duhamel, and O. Rioul, "Joint source-channel coding: Panaroma of methods," in *Proc. CNES Workshop Data Compression*, 1996, pp. 1232–1254.

[4] X. Wu, H. D. Mittelmann, X. Wang, and J. Wang, "On computation of performance bounds of optimal index assignment," *IEEE Trans. on Communications*, vol. 59, no. 12, pp. 3229–3233, Dec. 2011.

[5] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," in *Proc. 14th Int. Joint Conf. on Artificial Intelligence*, vol. 1, Montreal, CA, Aug. 1995, pp. 448–453.

[6] D. Lin, "An information-theoretic definition of similarity," in *Proc. 15th Int. Conf. on Machine Learning (ICML '98)*, San Francisco, 1998, pp. 296–304.

[7] J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," in *Proc. Int. Conf. on Research on Comp. Linguistics (ROCLING X)*, Taiwan, Sept. 1997, pp. 19–33.

[8] G. Miller, "Wordnet: An on-line lexical database," *Int. Journal of Lexicography*, vol. 3, no. 4, 1990.

[9] R. H. Mili, E. Bicknell, and M. Bletner, "Development and application of a metric on semantic nets," *IEEE Trans. on Systems, Man. and Cybernetics,*, vol. 19-33, no. 1, pp. 17–30, Nov. 1989.

[10] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences)*, 1st ed. W. H. Freeman, Jan. 1979.

[11] H. S. Witsenhausen, "The zero-error side information problem and chromatic numbers," *IEEE Trans. on Information Theory*, vol. 22, no. 5, pp. 592–593, Sept. 1976.

[12] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May. 1983.

[13] A. E. Gamal, L. A. Hemachandra, I. Shperling, and V. K. Wei, "Using simulated annealing to design good codes," *IEEE Trans. on Information Theory*, vol. 33, pp. 116–123, Jan. 1987.

[14] B. Hajek, "A tutorial on the survey of theory and applications of simulated annealing," in *Proc. IEEE Conf. on Decision and Control*, Ft. Lauderdale, FL., Dec. 1985, pp. 755–760.

[15] D. Hope, "Jws (java wordnet::similarity)," available online at http://www.sussex.ac.uk/Users/drh21/.