

УДК 621.396.669: 519.724

ИСПОЛЬЗОВАНИЕ ТЕОРЕТИКО-ИНФОРМАЦИОННЫХ ОЦЕНОК ПРИ ВЫБОРЕ СТРАТЕГИИ ПРОТИВОДЕЙСТВИЯ В УСЛОВИЯХ ИНФОРМАЦИОННОЙ ВОЙНЫ

В.Г. Хромых, А.В. Сычев

Воронежский государственный университет

В статье рассматривается возможность использования математического аппарата теории информации для выбора или построения более эффективных стратегий информационного противодействия в рамках информационной войны. На основе комбинаторного определения понятия количества информации (по Колмогорову) предложен подход к вычислению значимости элементов внутри сообщений, передаваемых по каналам связи. Использование распределения значимости элементов можно рассматривать как учет смысловых и лингвистических аспектов сообщений при выборе стратегии информационного противодействия. В качестве примера приведены оценки значимости для морфемного состава русского языка. Предложено перераспределять частотно-временной ресурс сигналов между элементами сообщения в зависимости от их значимости.

В последние годы значительно возросла роль телекоммуникационных средств в технологической, научной и социальной сферах жизни современного общества. Они являются основой глобальных и локальных информационных систем, влияющих на различные стороны жизни общества. Возможность достижения отдельными пользователями определенных преимуществ за счет специально выбранных воздействий на информационные потоки составляет сущность информационной войны [1].

В настоящее время наиболее полно исследованы различные аспекты информационных войн, связанных с конфликтом в спектре электромагнитных волн [2],[3], то есть фактически рассматривалась только часть процесса обмена информацией, а именно этап, на котором носителем информации является физический сигнал в канале связи. В некоторых случаях, например радиолокации, такой подход, по-видимому, является исчерпывающим. Однако, если по телекоммуникационным каналам передается семантическая информация, определенная на множестве лингвистических полей, то в случае сигнального подхода из рассмотрения выпадают основные этапы. Это, в первую очередь, этапы образования информации и ее воздействия на адресата. Целевое эффективное

воздействие на информационные потоки возможно только при учете этих этапов.

Трудности исследований моделей в общем виде заключаются в их математической формализации. Первым этапом такой формализации, по-видимому, должен быть этап интерпретации лингвистических и семантических аспектов сообщений в рамках формализма Шеннона-Колмогорова.

Конкретно, исследованию подлежит распределение элементов плана содержания (смысловых составляющих) внутри сигнального представления сообщения. Это распределение может быть использовано в качестве отправной точки для определения необходимого целенаправленного воздействия.

Общепринятый в теории информации вероятностный подход к измерению количества информации как раз и исходит из указанного выше представления сообщений. Однако, любое сообщение имеет ценность лишь в том случае, если оно имеет некоторый *смысл*, или, иначе говоря, ему может быть приписано некоторое *семантическое* значение. Формализованное представление этого аспекта дается в [4] следующим образом

Пусть B_1 - алфавит языка источника сообщений $L_1 = \langle \lambda_1, S_1 \rangle$, B_2 - алфавит языка адреса-

та $L_2 = \langle \lambda_2, S_2 \rangle$, $A = \{a_1, \dots, a_q\}$ - алфавит канала, $f: B_1^+ \rightarrow A^+$, $g: A^+ \rightarrow B_2^+$. Алфавит языка L_1 включает в себя λ_1 - собственно язык, то есть множество сообщений, и $S_1 \subseteq \lambda_1 \times \lambda_1$ - отношения синонимии в нем. Предполагается, что $\langle \alpha, \beta \rangle \in S_1$ означает, что информация в сообщениях α и β примерно одинакова. Если ввести функцию, задающую семантику сообщений $\psi = \psi(x)$, то из $\langle \alpha, \beta \rangle \in S_1$ следует, что $\psi(\alpha) = \psi(\beta)$. Главное требование, которое должно выполняться при передаче сообщений через канал связи, выражается условием согласованности:

$$\psi_1(\alpha) = \psi_2(g(f(\alpha))) \text{ для любого } \alpha \in \lambda_1.$$

Очевидно, что не все элементы, составляющие сообщение, равноценны с точки зрения формирования семантики, и учет этого обстоятельства позволяет построить более эффективный подход к распределению ограниченных ресурсов между ними в условиях информационной войны.

Исходя из вышесказанного, разработка любого алгоритма или системы передачи информации должна начинаться с выбора адекватной формы представления сообщений на основе предварительного анализа значимости с точки зрения семантики образующих эти сообщения элементов. При построении алфавита источника предпочтение должно отдаваться тем элементам, с которыми можно соотнести элементарные семантические значения.

Следующий этап - вычисление коэффициентов значимости символов алфавита источника. В качестве данной характеристики предлагается использовать *энтропию*.

В теории информации это понятие вводится следующим образом. Если известны априорные вероятности отдельных символов алфавита источника p_i , $i = 1, 2, \dots, N$, то энтропия источника определяется как $H_1 = \sum_i p_i \cdot \log_2 p_i$. Существенный недостаток такого определения заключается в использовании вероятностного подхода, который, как нам представляется, недостаточно адекватен при анализе смысловой структуры сообщений.

А.Н. Колмогоровым были предложены три подхода к определению количества информации: вероятностный, комбинаторный и алгоритмический [5]. Для анализа структуры системы, порождающей сообщения, удобно использовать комбинаторный подход. Идея комбинаторного подхода заключается в том, что при оценке количе-

ства информации, содержащейся в переменной $x \in X$ относительно связанной с ней переменной $y \in Y$, используется тот факт, что не все пары (x, y) , принадлежащие прямому произведению $X \times Y$, являются возможными. Условная энтропия определяется равенством $H(y|x) = \log_2 N(Y_a)$, где Y_a - множество возможных пар (a, y) , а $N(Y_a)$ - число элементов в этом множестве. Информация в x относительно y выражается формулой

$$I(x:y) = H(y) - H(y|x).$$

Введем многомерное пространство признаков C_1, C_2, \dots, C_d для символов алфавита источника. Тогда любой символ может быть представлен в виде точки из этого пространства. Если известны только k признаков символа, то им можно поставить в соответствие все точки из подпространства размерностью $d-k$ (остаточный класс), построенное из исходного пространства размерности d при фиксированных значениях C_1, C_2, \dots, C_k .

Для анализа структуры системы, порождающей сообщения, можно провести исследование зависимости усредненной по остаточным классам символов энтропии

$$\bar{H}(C_1, C_2, \dots, C_k) = \frac{1}{N} \sum_{C_1 \in C_1} \sum_{C_2 \in C_2} \dots \sum_{C_k \in C_k} N(C_1, C_2, \dots, C_k) \cdot \log_2 N(C_1, C_2, \dots, C_k)$$

$$N = \sum_{C_1} \sum_{C_2} \dots \sum_{C_k} N(C_1, C_2, \dots, C_k), \quad 1 \leq k \leq D$$

в зависимости от фиксированных значений C_1, C_2, \dots, C_k в многомерном пространстве признаков символов: C_1, C_2, \dots, C_D .

Для оценки значимости признака C_k можно использовать оценку

$$V(C_k) = \frac{\bar{H}(0) - \bar{H}(0, \dots, C_k, \dots, 0)}{\bar{H}(0)}.$$

В качестве примера рассмотрим текстовые сообщения на русском языке. В работах [6] и [7] предложен подход к кодированию этих сообщений, основанный на их поморфемном представлении, т.е текстовое сообщение представляется в виде последовательности приставок, корней, суффиксов и окончаний, а затем каждая из полученных морф кодируется определенным об-

разом. Для названных морфем выделены следующие признаки.

1. Приставка.

1) исходная часть речи; 2) результирующая часть речи; 3) словообразовательное значение; 4) сочетаемость; 5) валентность.

2. Суффикс.

Общие атрибуты:

1) исходная часть речи; 2) результирующая часть речи; 3) словообразовательное значение; 4) сочетаемость; 5) валентность.

В зависимости от **результатирующей части речи**:

Существительное.

а) род; б) одушевленность; в) число.

Глагол.

а) вид;

3. Окончание.

Общие атрибуты:

1) часть речи; 2) модель словоизменения; 3) число.

В зависимости от **части речи**:

Существительное.

а) падеж

Прилагательное.

а) род; б) краткость.

Глагол.

а) время; б) лицо; в) наклонение; г) залог.

4. Корень.

Для некоторых корней, особенно для тех, которые непосредственно сочетаются с окончаниями, можно было ввести грамматические признаки, однако данная работа проведена не была. Хотя можно было бы в качестве таких признаков фиксировать все возможные аффиксальные окружения для каждой корневой морфемы.

Ниже приводится таблица, в которой приводятся оценки значимости выделенных признаков морфем.

Если действия противоборствующей стороны строятся без учета семантических аспектов объекта-сообщения M , то эффективность этих действий, выражаемая некоторой величиной D , является функцией от энергии P^M сигнала-сообщения M и усилий, затрачиваемых противоборствующей стороной и выражаемых некоторой интегральной величиной P^C :

$$D = D(P^M, P^C).$$

Переход от чисто энергетического подхода в отношении стратегии противодействия к учету семантических аспектов, позволяет перераспределить фиксированный ресурс между элементами сообщения $M = (m_1, m_2, \dots, m_l)$ с учетом их значимости, которая может быть описана функцией $V = V(m_i)$. В этом случае

$$D(P^M, P_1^C) = D(P^M, P_2^C, V(m_i)), \quad P_1^C \leq P_2^C,$$

т.е. противоборствующей стороне при отсутствии информации о распределении ресурса в зависимости от значимости элементов $V(m_i)$ для достижения того же эффекта потребуется затратить больше усилий, чтобы оказать необходимое воздействие на сообщение. В качестве альтернативы эта сторона может провести информационную разведку и исходя из ее результатов выбрать подходящее воздействие:

$$P^C = P^C(V(m_i)).$$

Тип морфы	Общее число морф	Нулевая энтропия	1	2	3	4	5	6	7	8	9	10	11	12
Приставка	220.00	7.78	5.94	5.92	6.90	3.81	6.52							
Значимость %			23.66	23.92	11.33	51.04	16.21							
Суффикс	625.00	9.29	7.52	7.73	8.55	3.98	5.48	8.62	9.13	7.88	7.61	9.26	9.02	
Значимость %			19.03	16.77	7.94	57.15	41.00	7.19	1.70	15.16	18.06	0.30	2.88	
Окончание	330.00	8.37		7.00			3.67	7.33	5.57		6.91	8.15		7.85
Значимость %				7.73			56.13	12.39	33.42		17.41	2.59		6.17

В таблице перечислены следующие признаки:

1. Исходная часть речи; 2. [Результирующая] часть речи; 3. Валентность; 4. Словообразовательное значение; 5. Модель словоизменения (внутрисловный контекст); 6. Число; 7. Падеж; 8. Одушевленность; 9. Род; 10. Краткость; 11. Вид.

ЗАКЛЮЧЕНИЕ

Переход от чисто энергетического подхода к проблеме информационного противоборства к учету семантических аспектов сообщений, передаваемых по каналам связи, позволяет при фиксированном ресурсе либо повысить устойчивость сообщений к действиям противодействующей стороны либо усилить эффективность противодействия. Перераспределение основано на анализе значимости элементов, составляющих сообщение, и заключается в том, что величина ресурса, выделяемого для конкретного элемента, определяется его значимостью. После перераспределения при той же величине воздействия искажение наименее значимых элементов сообщения не должно уже существенно сказываться на возможности восстановить смысл сообщения, используя естественную избыточность сообщений и наиболее значимые элементы сообщения, которые при таком подходе будут искажаться с меньшей вероятностью.

И наоборот, зная характер распределения значимых элементов внутри сообщения противодействующей стороны, можно более эффективно распределить усилия, сосредоточив их на наиболее ключевых элементах, что позволяет нанести ей наибольшие потери.

Практическая реализация предложенного в работе подхода предполагает использование программно-аппаратных средств, что ввиду широкого спектра их возможностей не представляет собой серьезной технической проблемы.

СПИСОК ЛИТЕРАТУРЫ

1. Расторгуев С.П. Информационная война. - М. Радио и связь. 1998. - 415 с.
2. Палий А.И. Радиоэлектронная борьба. М.: Воениздат, 1989. - 350 с.
3. Радиотехника. Тематические выпуски «Конфликт в спектре электромагнитных волн».
4. Марков А.А. Введение в теорию кодирования. - М.: Наука, 1982. - 192 с.
5. Колмогоров А.Н. Теория информации и теория алгоритмов. - М.: Наука, 1987. - 304 с.
6. Оценка информационной емкости систем передачи информации, основанных на морфемном кодировании / Хромых В.Г., Сычев А.В. // Вестн. ВГУ. Сер 2, Естественные науки, 1996. №2. С. 137-144.
7. Хромых В.Г., Сычев А.В. Исследование корректирующей способности поморфемного кодирования текстовых сообщений, передаваемых через канал связи с шумом // Изв. высш. учеб. заведений. Радиотехника. 1998. №10. С. 35-42.