

УДК 004.822

## Семантические инфокоммуникации как следующий этап развития систем связи

Вершков П. А., Казачков В. О., Шевцов В. А.

**Постановка задачи:** в настоящее время в телекоммуникационных системах наблюдаются колоссальные темпы развития. При этом одной из причин является активное внедрение искусственных нейронных сетей и технологий глубокого обучения. Данная особенность позволяет реализовать принципиально новый подход на основе извлечения семантики в построении систем связи, применимых для различных областей от гражданской до военной направленности, делая акцент на основном смысле передаваемой информации в отличие от традиционной формы передачи бинарного потока, характерного системам предыдущих поколений. Задачей работы является исследование семантических инфокоммуникаций. **Целью работы** является формирование аналитического обзора на основе агрегации и структуризации информации о семантических инфокоммуникационных системах от зарождения концепции и ее основ до современных реализаций разнообразных подходов в зависимости от типа передаваемой информации на прикладном уровне. **Используемые методы:** реализация семантических инфокоммуникаций базируется на аппарате высшей математики. В частности, линейная алгебра является основным инструментом математических вычислений, необходимых для реализации искусственной нейронной сети. Нейронная сеть отвечает за извлечение семантики из передаваемых данных и является ключевым элементом обработки сигнала в структуре семантической инфокоммуникационной системы. **Результаты и новизна:** элементом практической новизны работы являются систематизированные особенности построения системы связи с семантическим уровнем обработки информации. В частности, приведена общая структурная схема, рассмотрены основные подходы реализации для различных типов данных, таких как аудио, текстовые и графические данные. Основной акцент сделан на детальную проработку случая передачи и приема графических данных. Выведена структура и типовой подход извлечения семантики на основе сверточной нейронной сети для обработки изображений. Структурированы основные показатели помехоустойчивости. **Практическая значимость:** представленная информация будет полезна в научной и инженерной практике при исследовании и разработке семантических инфокоммуникаций.

**Ключевые слова:** семантические инфокоммуникации, семантическое кодирование, гетерогенные системы связи, мобильная связь, архитектура системы связи, нейронные сети, семантическое взаимодействие, сверточные нейронные сети.

### Введение

В последние годы наблюдается стремительное развитие инфокоммуникационных технологий, продиктованное целым рядом разнородных причин, например, колоссальными темпами роста мобильного трафика, разнообразием

#### Библиографическая ссылка на статью:

Вершков П. А., Казачков В. О., Шевцов В. А. Семантические инфокоммуникации как следующий этап развития систем связи // Системы управления, связи и безопасности. 2023. № 3. С. 267-292. DOI: DOI: 10.24412/2410-9916-2023-3-267-292

#### Reference for citation:

Vershkov P. A., Kazachkov V. O., Shevtsov V. A. Semantic infocommunications as the next stage in the development of communication. *Systems of Control, Communication and Security*, 2023, no. 3, pp. 267-292 (in Russian). DOI: 10.24412/2410-9916-2023-3-267-292

траффика с учетом специфики каждой отдельно взятой системы и т.д. К новым поколениям систем связи предъявляются все более жесткие требования, что неизбежно приводит к разработке и применению новых подходов и методов обработки сигналов. Разнообразие инфокоммуникационных систем и устройств также неизбежно порождает необходимость решения задачи взаимодействия между ними. В условиях разнообразных сфер деятельности, таких как гражданские, военные, медицинские, промышленные и другие, сети функционируют в гетерогенной среде с разнообразными протоколами, структурами и форматами данных. Обеспечение совместной работы и обмена информацией между такими гетерогенными сетями представляет сложную задачу, которая требует инновационных подходов и решений. С развитием и применением глубокого обучения и искусственных нейронных сетей появилась возможность их использования в инфокоммуникациях для семантической обработки информации.

Концепция семантической инфокоммуникации в системах связи представляет собой акцент не только на традиционной передаче данных, а на достоверном кодировании и представлении информации с учетом ее семантического содержания и понимании передаваемой информации на приемной стороне, что позволяет создавать более эффективные системы связи. Несмотря на основное гражданское применение данного направления исследований на сегодняшний день, потенциал концепции семантических инфокоммуникаций позволяет в будущем использовать ее и под нужды систем связи специального назначения.

В данном аналитическом обзоре будет рассмотрена концепция семантических инфокоммуникаций от истоков ее возникновения до современных исследований в данной области в контексте потенциального применения в системах связи будущих поколений с углубленным рассмотрением особенностей передачи изображений.

## 1. Развитие концепции семантической инфокоммуникации

Слово «семантика» (значение) происходит от языков (естественных или формальных) и от концепции композиционности, согласно которой значение предложения определяется тремя составляющими:

- правило составления предложений (синтаксис);
- значение (семантика) каждого компонента;
- контекст.

Изучением семантики занимались многие лингвисты начиная с конца XIX века. Первым термин «семантика» ввел французский ученый М. Бреаль, однако, он рассматривал семантику как изменение значения слов с течением времени. Далее в 1938 г. Ч. Моррис [1], в рамках работы о науке семиотике, определил семантику как связь между символами, их значениями и объектами, которые могут ими обозначаться.

К. Шеннон [2] представил математический подход к описанию передачи информации и обмену сообщениями между двумя или более точками в системе связи. Шеннон ввел понятие информационной энтропии, которая описывает количество информации и связана с вероятностью появления определенных

символов в сообщении. Он также ввел понятие канала связи и предложил математический метод измерения пропускной способности канала связи. В статье Шеннон также представил концепцию шума, который может влиять на передачу сообщений, и рассмотрел методы кодирования сообщений для более эффективной передачи по каналу связи. В целом, данная работа принесла революционные идеи в области теории информации и стала основой для развития современных методов передачи информации и коммуникаций.

Продолжением исследований инфокоммуникаций стала статья [3], выпущенная в 1949 г. У. Уивером, которая является продолжением идей, изложенных в статье К. Шеннона. В своей части Уивер приводит 3 проблемы коммуникации:

- уровень А. Насколько точно могут быть переданы символы? (техническая проблема);
- уровень В. Насколько точно передаваемые символы передают желаемый смысл? (семантическая проблема);
- уровень С. Насколько эффективно полученное значение влияет на поведение желаемым образом? (проблема эффективности).

Математическая теория К. Шеннона описывает уровень А, на котором акцент делается на коде и канале передачи информации. В этой теории объем информации, содержащейся в сообщении, определяется множеством всех возможных сообщений и их вероятностями, независимо от их смыслового содержания. Шеннон предоставил строгое и формальное решение технической проблемы, заложив основы того, что сегодня известно как теория информации. Уивер считал, что математическая теория является чрезвычайно общей по своему охвату и рассматривает фундаментальные проблемы, но не учитывает тип данных, которые будут передаваться. При передаче сообщений возникают технические трудности, связанные с передачей символов, сигналов и изображений, а также семантические трудности, связанные с пониманием смысла сообщений получателем. Уивер считал, что схема системы связи, предложенная Шенноном, является достаточной для решения технической проблемы (уровень А), но при рассмотрении следующих уровней потребует дополнений или модернизаций. Например, в виде добавления между приемником и получателем еще одного блока – семантического приемника, в котором бы учитывались семантические характеристика сообщения. Уивер также расширил понятие «шума» на другие формы помех, такие как культурные, религиозные, нравственные различия, а также контекст, которые могут влиять на понимание сообщения. Он предложил концепцию «семантического шума», который возникает, когда человек интерпретирует сообщение не так, как задумывалось отправителем. Одним из важных вкладов Уивера была его идея о взаимосвязи между информацией и энтропией в системе коммуникации. Он показал, что информация не только передается, но и создается в процессе коммуникации, и что она может быть измерена через количество новых знаний, полученных реципиентом.

После У. Уивера исследование концепций, связанных с семантическим аспектом передачи информации, продолжали многие ученые. В 1952 г. Р. Карнап и Й. Бар-Хиллел [4] представляют теоретическую модель семантиче-

ской информации, определяют основные понятия и методы анализа, и обсуждают возможности применения этой модели в различных областях, включая науку, технику, общество и культуру. Статья начинается с определения основных понятий: «информация», «смысл» и «символ». Карнап и Бар-Хиллель утверждают, что информация связана со знаниями и смыслом, а не просто с данными или фактами. Они вводят понятие «символьной структуры», которая состоит из символов и отношений между ними, и говорят о том, что смысл информации зависит от этой структуры. Карнап предлагает использовать понятие «тестового элемента» для определения содержательной информации в сообщении. Тестовый элемент – это элемент, который может быть использован для определения, истинно ли данное утверждение или нет. Карнап также описывает несколько принципов, которые могут быть использованы для определения, какие утверждения имеют содержательную информацию, а какие – нет. Далее авторы представляют свою теорию о том, как информация может быть передана от отправителя к получателю. Они вводят понятие «языка» и утверждают, что для передачи информации необходимо использовать общий язык, который должен быть определен заранее, чтобы исключить проблему «неопределенности», которая возникает, когда два или более сообщений могут быть интерпретированы как содержащие одинаковую информацию. Они также говорят о том, что передача информации зависит от контекста и культурных условий, и что эти факторы необходимо учитывать при разработке моделей передачи информации. В заключение авторы обсуждают возможности применения своей модели в различных областях, включая науку, технику, общество и культуру. Они говорят о том, что понимание смысла и передача информации являются ключевыми факторами для развития человеческого знания и прогресса.

Первые исследования в области семантической инфокоммуникации носили теоретический характер, без явного описания какими методами и средствами необходимо решить данную проблему. Л. Клейнрок, один из основателей теории сетей и интернета, упоминал о семантическом кодировании в нескольких своих статьях, посвященных проблемам передачи информации в сетях. В статье [5] упоминается семантическое кодирование в контексте исследования процесса передачи сообщений по сети. Клейнрок описывает возможность использования семантического кодирования для уменьшения нагрузки на канал связи путем сжатия информации до ее смыслового содержания. Кроме того, автор отмечает, что использование семантического кодирования может помочь улучшить эффективность передачи информации в сети, особенно в условиях ограниченной пропускной способности канала связи. Таким образом, он признавал важность семантического кодирования для эффективной передачи информации в сетях и активно исследовал его возможности в своих работах. В начале 1960-х годов в СССР ученый Ю.А. Шрейдер, занимавшийся задачами оптимизации управления, решил разработать методы определения того, как характер и структура информации, содержащейся в объекте, влияют на ее прием и эффективность обработки для использования. В результате появилась семантическая теория информации Шрейдера, которая отличалась от теории информации Шеннона своими исходными положениями. В теории Шрейдера основ-



ной акцент делался на свойствах системы-приемника, воспринимающей и накапливающей информацию, а также на оценке ее семантического значения. При этом вопрос о канале передачи информации отодвигался на второй план. Основная идея семантической теории информации заключается в том, что смысловую информацию, воспринимаемую данной системой, можно оценивать по степени изменения содержащейся в системе собственной семантической информации за счет накопления внешней информации. Для того чтобы воспринимать внешнюю информацию, система-приемник должна обладать неким минимальным запасом знаний, который Шрейдер назвал «тезаурусом» [6]. Тезаурус – это структурированный словарь, представляющий собой сеть связанных между собой понятий и терминов. Он используется для описания смысловых отношений между словами и понятиями, а также для повышения точности поиска и извлечения информации. Шрейдер в своих работах акцентировал внимание на необходимости использования тезауруса для создания компьютерных систем, которые могли бы эффективно обрабатывать естественный язык. Он также разработал методику создания тезаурусов, основанную на анализе частотности слов и построении иерархических структур.

В иерархии интеллектуальной связи канал передачи использует основанное на теореме Шеннона внутреннее канальное кодирование и внешний семантический канал. В этом семантическом канале коды источника извлекают интересующую семантику, отличную от простого сжатия семантики, в сторону информационной энтропии.

Исследования в области семантического кодирования в компьютерных сетях начались в 1980-х годах. В телекоммуникациях семантическое кодирование было предложено как один из способов повышения эффективности передачи информации. Первоначально, семантическое кодирование было разработано для улучшения качества передачи аудио-сигналов, например, при передаче речи по телефонной линии. В [7] описывается алгоритм обнаружения слов в разговорных речевых сигналах низкого качества, например, записанных по телефону. Авторы предлагают новый подход, который комбинирует как синтаксические, так и семантические ограничения для повышения точности распознавания слов в речевых сигналах. Алгоритм использует языковые модели и синтаксический анализ для улучшения качества распознавания слов, а также учитывает контекстуальную информацию, чтобы улучшить точность распознавания слов, которые трудно различить на фоне шума и других искажений звука. Синтаксические ограничения включают в себя контекстную информацию, такую как части речи, порядок слов и грамматические правила, чтобы улучшить распознавание слов в предложении. Семантические ограничения включают в себя знание о том, какие слова могут и не могут находиться вместе в предложении, чтобы улучшить точность распознавания смысла целых предложений. Алгоритм работает следующим образом: входной сигнал речи разбивается на небольшие сегменты, называемые «акустическими векторами». Затем для каждого слова в словаре выполняется сравнение со всеми возможными комбинациями акустических векторов, чтобы найти наиболее вероятную последовательность векторов для данного слова. Затем с помощью синтаксических и семан-

тических ограничений ищутся наиболее вероятные комбинации слов в предложении.

В дальнейшем, с развитием цифровых технологий, семантическое кодирование стало использоваться для передачи цифровых данных, таких как изображения, видео и текстовые данные. Г. Хинтон – канадский ученый, который совместно с коллегами в 2006 г. предложил метод Deep Belief Networks (глубоких сетей вероятностных моделей), использующих семантическое кодирование для обучения нейронных сетей. Статья [8] исследует эту задачу с использованием семантических кодов (semantic output codes). Конечная цель обучения с учителем (supervised learning) заключается в том, чтобы обучить модель, которая может правильно классифицировать объекты на основе известных ранее примеров. Однако, в реальном мире мы не всегда можем иметь доступ к представительным примерам всех возможных классов, что делает задачу обучения с учителем более сложной. Решением проблемы является обучения без примеров (zero-shot learning), когда модель должна классифицировать объекты на основе знаний о свойствах классов, не участвовавших в обучении. В статье авторы предлагают новый метод, который использует семантические коды для классификации объектов. Они предлагают использовать условные случайные поля (conditional random fields, CRF) для моделирования связей между признаками и классами. В результате экспериментов на нескольких наборах данных, авторы показывают, что предложенный метод превосходит существующие методы zero-shot learning и достигает точности, сравнимой с методами обучения с учителем.

Одним из ключевых моментов в развитии семантического кодирования стало создание в 2013 г. нейронной сети Word2Vec, разработанной группой исследователей в Google. Эта модель позволяет представлять слова в виде векторов в многомерном пространстве, где близость векторов соответствует семантической близости слов.

Интересно отметить, что в 2012 г. в эталонной модели интероперабельности в соответствии с ГОСТ Р 55062-2012 также вводится семантический уровень. Этот уровень дополняет стандартные процессы передачи данных, как знаково-символьно-сигнальной формы представления информации, еще и рассмотрением ее смыслового содержания. Интероперабельность – способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена [9]. Теория интероперабельности имеет ряд точек корреляции с концепцией семантической инфокоммуникации. Так отмеченная ранее в [3-4] зависимость смысла передаваемой информации от контекста – это одна из основных проблем обеспечения интероперабельности на семантическом уровне. А введенный в [6] термин «тезаурус» в теории интероперабельности называется «предметной областью». Эталонная модель интероперабельности (рис. 1) кроме семантического уровня содержит также организационный и технический.

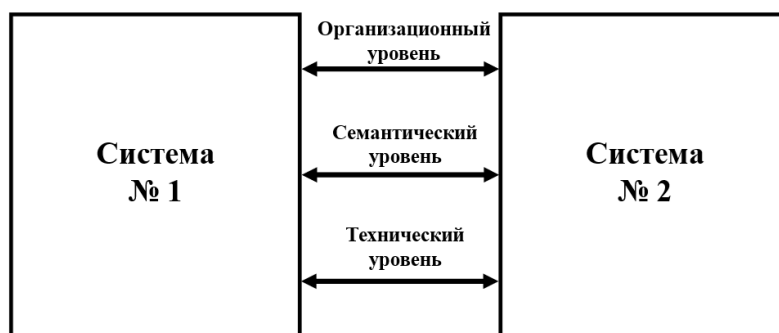


Рис. 1. Эталонная модель интероперабельности

Технический уровень описывает синтаксис или форматы передаваемой информации. Семантический уровень описывает содержательную сторону обмениваемой информации. Семантическая интероперабельность относится к способности информационных систем и устройств понимать и интерпретировать смысл передаваемой информации. Это понятие выходит за рамки простого обмена данными на техническом уровне, так как фокусируется на обеспечении информационного взаимодействия между системами на уровне смысла информации и контекста. Основной задачей семантической интероперабельности является обеспечение согласованности смысловых аспектов передаваемой информации между различными информационными системами. Это достигается путем использования совместных предметных областей, целей взаимодействия, контекстов, а также минимизации терминов, имеющих неоднозначное толкование с учетом национальных, религиозных, гендерных и др. различий. Статья [10] рассматривает создание концепции семантической интероперабельности и предоставляет описание семантического взаимодействия. В ней также представлена классификация агентов, которые взаимодействуют на семантическом уровне, и особенности их коммуникации в зависимости от категории. В работе [11] более детально описан процесс обмена информацией между агентами на семантическом уровне. Для успешного взаимодействия между агентами (людьми и техническими системами) необходим обмен сообщениями. Если агенты могут интерпретировать смысл информации, ее семантическое содержание, то их способность понимать друг друга улучшается. Таким образом, концепция семантической инфокоммуникации может послужить развитием «технического подхода» к информационному взаимодействию с последующим переходом на «семантический уровень» с целью повышения достоверности, полноты и контекста информационных сообщений.

Ученые, работы которых рассмотрены ранее и многие другие внесли важный вклад в развитие семантической инфокоммуникации, а их работы стали отправной точкой для многих современных исследований в этой области.

## 2. Применение семантических инфокоммуникаций

Семантические инфокоммуникации (semantic communications, SC) — это концепция коммуникации, который извлекает семантическую информацию из

источника и кодирует ее для передачи в канале с шумом. В отличие от традиционных коммуникаций, SC не требует, чтобы декодированная последовательность на стороне приемника строго соответствовала закодированной последовательности на стороне отправителя. Требуется только, чтобы восстановленная на стороне приемника семантическая информация (semantic information, SI) соответствовала переданной семантической информации на стороне отправителя, тем самым переходя от традиционной безошибочной передачи битов к лаконичной семантической передаче. Традиционная связь является синтаксически закодированной передачей исходного содержимого на основе символов с акцентом на безошибочные символы. SC — это извлечение, кодирование и передача характеристик/смысла исходного содержимого с акцентом на поддержание согласованности полученной информации со смыслом исходного содержимого. SC преодолевает традиционную передачу на уровне битов для реализации передачи на семантическом уровне, в следствие чего меняется и архитектура системы связи. На рис. 2 приведена структурная схема системы SC. По сравнению с архитектурой традиционных инфокоммуникаций, у SC есть дополнительный слой — семантический слой (обозначен оранжевым цветом).

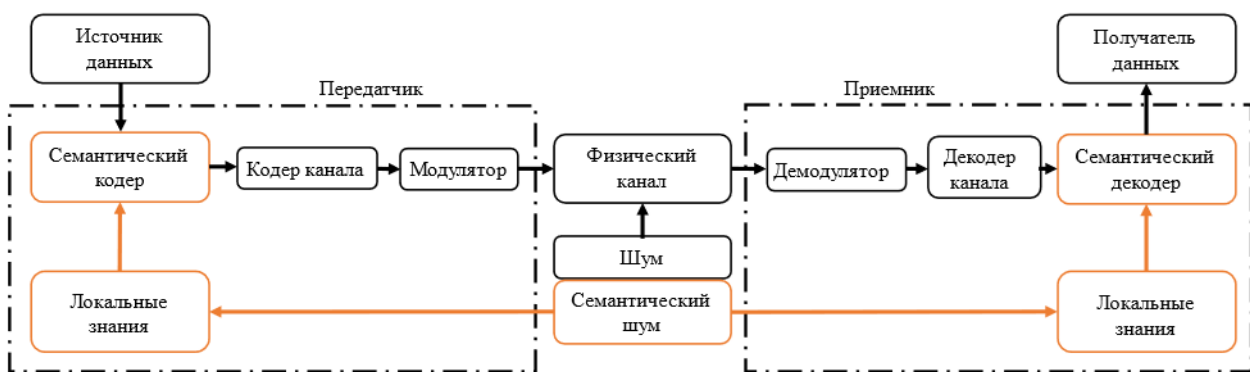


Рис. 2. Структурная схема системы связи семантической инфокоммуникации

Семантический слой реализует извлечение семантических особенностей исходной информации и семантическое восстановление полученной информации для осуществления эффективной передачи и коммуникации. Также важным аспектом семантической системы является то, что для эффективной передачи необходимо обмениваться локальными знаниями коммуницирующих сторон, чтобы обеспечить соответствие процессов понимания и вывода для всех исходных данных. Несоответствие локальных знаний приводит к семантическому шуму и неоднозначности, даже при отсутствии синтаксических ошибок в физической передаче. Семантический шум относится к любым искажениям, ошибкам или потерям смысла, которые могут возникнуть в процессе передачи семантической информации. Физический шум, связан с помехами в канале связи, семантический шум так же связан с несоответствием или потерей смысла передаваемой информации. Семантический шум может возникать из-за различных факторов, в том числе таких как различия в понимании терминов, контекстуальные несоответствия, неправильная интерпретация или неправильное кодирование и декодирование семантической информации.



При переходе к семантическим инфокоммуникациям стоит обратить внимание на работы Карнапа и Бар-Хиллеля [4], которые предложили концепцию семантической энтропии, используя логическую, а не статистическую вероятность для измерения семантической энтропии предложения. Логическая вероятность предложения измеряется вероятностью того, что предложение истинно во всех возможных ситуациях [12]. Тогда семантическая информация сообщения  $e$  определяется следующим образом:

$$H_S(e) = -\log_2(m(e)), \quad (1)$$

где:  $m(e)$  – логическая вероятность  $e$ .

Однако этот показатель привел к парадоксу, заключающемуся в том, что любой факт имеет бесконечное количество информации, когда он противоречит сам себе, т. е.  $H_S(e \wedge \neg e) = \infty$ .

Флориди [13] разрешил парадокс в предложении Карнапа и Бар-Хиллеля [4], приняв относительное расстояние семантики для измерения количества информации. Бао и др. [12] определили семантическую энтропию сообщения  $x$  следующим образом:

$$H_S(x) = -\log_2(m(x)), \quad (2)$$

где логическая вероятность  $x$  определяется выражением:

$$m(x) = \frac{\mu(W_x)}{\mu(W)} = \frac{\sum_{w \in W, w|=x} \mu(w)}{\sum_{w \in W} \mu(w)}, \quad (3)$$

где:  $W$  – набор символов источника;  $|=$  – отношение удовлетворения соотношений;  $W_x$  – множество моделей для  $x$ ;  $\mu$  – вероятностная мера.

Помимо логической вероятности, есть несколько определений семантической энтропии, основанных на разных предпосылках [14-15]. Д'Альфонсо [16] использовал понятие правдоподобия для количественной оценки семантической информации. Колчинский и Вольперт [17] определили семантическую энтропию как синтаксическую информацию, которой обладает физическая система о своей среде и которая необходима системе для поддержания собственного существования.

В целом, характеристики SC можно резюмировать следующим образом:

- передача информации носит избирательный характер. В частности, семантическая информация понимается до ее передачи. Полезная информация извлекается и передается, а остальная информация сжимается, чтобы уменьшить использование полосы пропускания;
- передаваемая информация в SC ориентирована в основном на задачи, а не на информацию. В первую очередь необходимо передать смысл, а не точную формулировку;
- эффективность SC обычно измеряется семантическим сходством, а также варьируется в зависимости от различных типов передаваемого контента.

Для применения семантической обработки данных необходимо рассмотреть 3 аспекта:

- 1) как выделить семантическую информацию;
- 2) как передать семантическую информацию;
- 3) как оценить семантическую информацию.

Для извлечения смысла в концепции семантических инфокоммуникаций в модели связи применяется *semantic extraction (SE)*, что позволяет передавать только ту информацию, которая является релевантной для получателя, исключая необработанные данные [11]. SE является переходом от концепции «передачи перед пониманием» к парадигме «понимания перед передачей». В SE для семантико-ориентированных инфокоммуникаций можно выделить 2 основных подхода: глубокое обучение (*deep learning, DL*) и обучение с подкреплением (*reinforcement learning, RL*).

DL является подтипом машинного обучения, который ориентирован на обучении нейронных сетей с большим количеством слоев. DL основан на идее построения иерархических представлений данных, которые позволяют извлекать семантику из входных данных. Глубокое обучение обычно применяется при решении задач обработки изображений, распознаванию речи, обработке естественного языка и другим задачам, где требуется анализ больших объемов данных.

RL — это подход к машинному обучению, который связан с изучением оптимального поведения агента в определенной среде. В контексте обучения с подкреплением термин «агент» относится к компоненту, который выполняет действия и принимает решения в среде. Агент является центральным элементом в RL и представляет собой автономную систему, которая взаимодействует с окружающей средой, получает обратную связь в виде «награды» или «штрафа» за свои действия и стремится максимизировать накопленную суммарную награду. Агент в RL взаимодействует с окружающей средой и может наблюдать ее состояние, принимает решения и выполняет определенные действия в среде на основе своих наблюдений, получает обратную связь в виде награды или штрафа от среды за каждое выполненное действие или состояние. Агенты в RL могут быть реализованы с использованием различных методов, например, эволюционных алгоритмов. Данные методы применяются во многих областях, включая робототехнику, игровую индустрию, финансовый анализ и другие, где необходимо научиться принимать оптимальные решения в изменяющейся среде на основе накопленного опыта.

Так как в зависимости от задачи могут применяться различные варианты реализации концепции семантической инфокоммуникации, далее будут приведены примеры SE для различных типов данных.

### 3. Особенности семантических инфокоммуникаций для передачи различных типов данных

#### 3.1. Извлечение семантики из аудиоданных

В [18] исследуется применение модели SemCom для аудиосигналов. Авторы разработали аудио SE на основе модели Wav2Vec, основанной на глубоком обучении и обработке естественного языка (natural language processing, NLP). Семантический кодер состоит из двух связанных сверточных нейронных сетей (convolutional neural network, CNN): FE и feature aggregator (FA). FE извлекает первичные аудио-функции из исходного вектора аудио, а FA объединяет их в более высокоуровневую скрытую переменную, содержащую семантические связи между контекстуальными аудио-функциями. Семантический декодер также основан на архитектуре Wav2Vec и состоит из двух симметричных CNN, называемых «feature decomposer» (FD) и «audio generator» (AG). Однако у этой модели есть ограничения в извлекаемой семантической информации, и ее производительность зависит от фиксированных условий канала, что ограничивает ее применение в канале с помехами.

В [19] предложена структура DeepSC-S, система передачи речи на основе глубокого обучения, спроектированная как единый канал связи с добавлением семантического кодера и семантического декодера. CNN основана на концепции «внимания» (attention). Функцию внимания можно описать как отображение запроса и набора пар «ключ-значение» в выходные данные, где запрос, ключи, значения и выходные данные являются векторами. Выходные данные вычисляются как взвешенная сумма значений, где вес, присвоенный каждому значению, вычисляется с помощью функции совместимости запроса с соответствующим ключом. Основанная на «внимании» двумерная (2D) сверточная нейронная сеть разработана в семантическом кодере и декодере для восстановления переданных речевых выборок путем обучения и извлечения речевых сигналов. Передатчик состоит из каскада семантического кодера и канального кодера, извлекающего семантические особенности речи отдельными нейронными сетями. Приемник состоит из семантического декодера и канального декодера, преобразующего изученные и извлеченные семантические особенности речи в сигналы. Преимущество заключается в том, что семантический кодер/декодер и канальный кодер/декодер работают совместно (joint source-channel coding, JSCC), что обеспечивает высокую надежность и адаптивность.

#### 3.2. Извлечение семантики из текстовых данных

Под влиянием успеха глубокого обучения в области обработки естественного языка, в частности в машинном переводе, авторы в работе [20] представляют первоначальную реализацию семантических инфокоммуникаций для передачи текста. В их исследовании рассматривается простая модель системы, в которой передатчик отправляет предложения получателю, используя ограниченное количество бит через канал с потерями. В предложенной схеме слова сначала представляются вектором вложения с использованием GloVe (global vectors for word representational), который является предварительно обученной

таблицей поиска, используемой для извлечения семантической информации. Затем используются кодер и декодер на основе долгой краткосрочной памяти (long short-term memory, LSTM), в которых вектор вложения предыдущего определенного слова берется в качестве входных данных для следующего шага, а алгоритм поиска по графу используется для нахождения наиболее вероятных последовательностей слов [21]. В этом смысле SI может быть встроена в восстановление предложений.

В [22] исследователи предложили архитектуры SC, основанные на глубоком обучении. Архитектуры разрабатываются в соответствии с типами передаваемой информации, в работе рассматривается передача текста на основе глубокого обучения. Семантический уровень адресует обработку семантической информации для кодирования и декодирования с целью извлечения семантической информации, уровень передачи гарантирует, что семантическая информация может корректно передаваться по каналу связи.

### 3.3. Извлечение семантики из графических данных

Работа [23] посвящена на передаче изображений, где устройство интернета вещей (internet of things, IoT) передает изображения на сервер для выполнения распознавания. Устройство IoT поддерживает прямую связь «точка-точка» с сервером. В отличие от обычных моделей коммуникации, где используются несколько модулей каскадного типа, авторы предлагают схему совместной передачи и распознавания смысла, построенную с использованием глубокого обучения. Авторы разделили глубокую нейронную сеть (deep neural network, DNN) на две части: первые слои выполняют функцию извлечения признаков на передатчике, а остальные слои служат распознавателем на приемнике. Работа [24] включает в модель SE механизм извлечения признаков (FE – feature extraction), который широко используется в компьютерном зрении, для адаптивного кодирования семантики и канала. В предложенной схеме совместного кодирования семантики и канала одна нейронная сеть выполняет функции извлечения признаков FE и модуля внимания (attention feature, AF [25-26]). Предложенная схема показывает высокую устойчивость, универсальность и адаптивность к широкому диапазону отношения сигнал-шум (signal-to-noise ratio, SNR).

Как следует из представленной информации, семантическое кодирование может быть успешно применено для передачи различных типов данных (текст, звук, изображения). При этом существует ряд факторов, которые подтверждают преимущества передачи именно в контексте графической информации:

- эффективная визуализация сложных концепций. Графическая информация, такая как диаграммы, графики и схемы, способна наглядно представить сложные концепции, взаимосвязи и важные аспекты данных. Это позволяет легче и понятнее воспринимать информацию и улучшает общее понимание;
- универсальность и доступность. Визуальные элементы являются универсальным средством коммуникации, которое понятно людям разных

- культур и языков. Они не требуют особых навыков или специфического языка и доступны для широкой аудитории;
- эффективная передача информации. Графическая информация может быть компактной и содержательной. Она может сократить объем информации, исключая излишние детали и фокусируясь на ключевых аспектах. Такой подход позволяет более эффективно передавать информацию и улучшить ясность коммуникации.

Исходя из вышеперечисленных преимуществ в следующем разделе приведено более детальное рассмотрение семантических инфокоммуникаций и семантического кодирования для передачи именно графической информации.

#### **4. Семантические инфокоммуникации для передачи графической информации**

Применение семантической информации для передачи изображений представляет собой актуальную исследовательскую область в контексте передовых технологий коммуникации и может охватывать разнообразные области, включая медицинскую диагностику, системы видеонаблюдения, мониторинг объектов, анализ критических ситуаций, виртуальную и дополненную реальность, сферы искусства и дизайна. Огромный потенциал содержится и для применения семантической информации для передачи изображений в военной сфере для улучшения ситуационной осведомленности, принятия решений и координации операций. Извлечение семантики позволяет эффективно обрабатывать большие объемы разведывательных данных, устанавливать связи между событиями, аннотировать данные метаданными для более точной интерпретации, обеспечивать интероперабельность разнородных систем и т.д.

Семантическое кодирование и передача изображений основаны на использовании семантических моделей, которые позволяют более эффективно передавать и интерпретировать содержание изображений между отправителем и получателем. Традиционные методы передачи изображений, основанные на сжатии и кодировании, обычно не учитывают семантику содержания. Вместо этого, они сконцентрированы на сохранении структурных и визуальных аспектов графического формата данных. В процессе семантического кодирования изображений, перед отправкой, происходит анализ содержания с помощью семантических моделей. Семантический анализ позволяет выделять ключевые семантические признаки и описание содержания изображения, что делает его более понятным для получателя.

На рис. 3 представлена одна из наиболее распространенных архитектур для передачи изображений в семантических инфокоммуникациях. В данных архитектурах используются сверточные нейронные сети выполняющие совместно функции семантического кодера и кодера канала, либо кодер канала может быть представлен в классическом виде. Сверточные нейронные сети эффективно работают с изображениями благодаря особенностям их архитектуры, которая учитывает специфику такого типа данных. В отличие от обычных нейронных сетей, сверточные нейронные сети организованы в слоях, где нейроны рас-



полагаются в трех измерениях: ширина, высота и глубина. Специфика трехмерной организации слоев в сверточных нейронных сетях позволяет им лучше улавливать локальные зависимости между пикселями изображения и выявлять важные признаки на разных уровнях детализации. Каждый нейрон в сверточном слое соединен только с небольшой локальной областью предыдущего слоя, что позволяет им эффективно обрабатывать изображения с большим количеством пикселей.

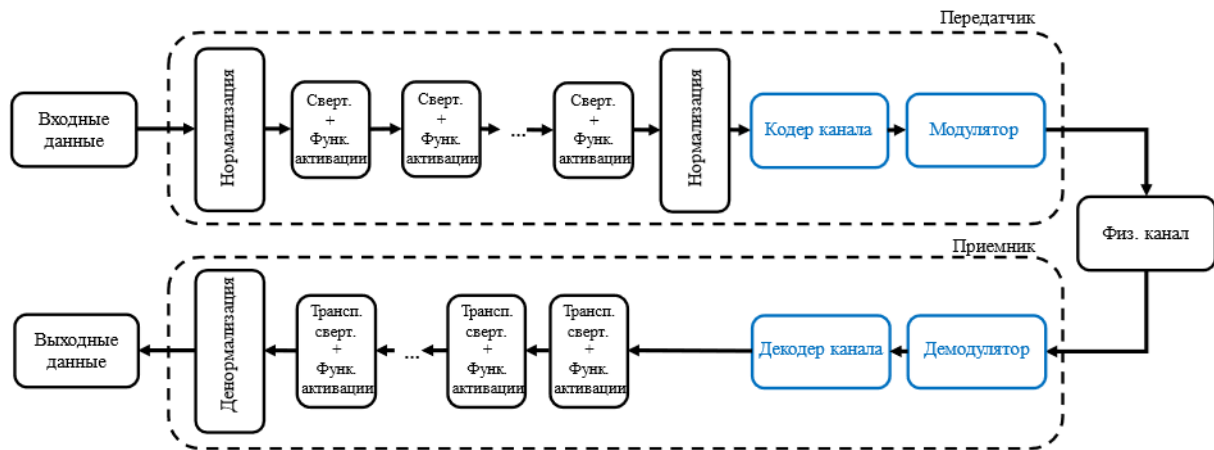


Рис. 3. Архитектура SC для передачи графических данных

В данных архитектурах изображение подается на входной слой сети. Обычно это матрица пикселей, где каждый пиксель представляет интенсивность или цвет соответствующего пикселя изображения. Слой нормализации является важным компонентом сверточных нейронных сетей и предназначен для улучшения стабильности и скорости обучения модели. Он выполняет нормализацию активаций на выходе предыдущего слоя. Существует несколько различных типов слоев нормализации, включая Batch Normalization (BN), Layer Normalization (LN) и Instance Normalization (IN). Различные типы слоев нормализации используются в разных контекстах и имеют свои особенности. BN является одним из наиболее широко используемых типов слоев нормализации. Он выполняет нормализацию активаций по мини-пакетам (batch) данных. Для каждого признака входных данных BN вычисляет среднее и стандартное отклонение по всему мини-пакету. Layer Normalization нормализует активации на уровне каждого отдельного слоя, а не по мини-пакетам. LN обычно используется в моделях, где размеры мини-пакетов могут быть непредсказуемыми или где batch-нормализация не применима. Instance Normalization нормализует активации на уровне каждого отдельного примера (instance) в мини-пакете данных. IN обычно используется в задачах, где требуется нормализация активаций для каждого примера независимо от остальных [27].

Сверточный слой является основным компонентом сверточных нейронных сетей, которые применяются для обработки данных с сетчатой структурой, таких как изображения. Он выполняет операцию свертки, которая включает применение фильтров к входным данным для извлечения признаков и выявления пространственных шаблонов. Сверточный слой состоит из набора обучае-

мых фильтров, также известных как ядра или детекторы признаков. Каждый фильтр представляет собой матрицу небольшого размера (обычно 3 на 3 или 5 на 5 [28-30]), которая «скользит» или свертывается по входным данным. Фильтр применяется к каждой возможной позиции входных данных, вычисляя скалярное произведение между весами фильтра и соответствующими значениями входа (рис. 4а). Во время операции свертки фильтр перемещается с фиксированным шагом (обычно 1 пиксель для более точной классификации [31]) по входным данным. На каждой позиции вычисляется скалярное произведение путем поэлементного умножения значений фильтра и соответствующих значений входа, за которым следует суммирование результатов. Эта операция порождает одно значение, которое представляет ответ или активацию фильтра в данной позиции. Операция свертки повторяется для каждого фильтра в сверточном слое. В результате генерируется несколько карт признаков, где каждая карта признаков соответствует определенному фильтру. Эти карты захватывают различные аспекты или шаблоны входных данных, такие как края, текстуры или более высокоуровневые признаки. Этот иерархический процесс извлечения признаков является основой успеха сверточных нейронных сетей в различных задачах, таких как классификация изображений, обнаружение объектов и сегментация изображений.

Функция активации в сверточных нейронных сетях играет важную роль в введении нелинейности в модель. Активационная функция применяется к каждому выходу или активации сверточного слоя, преобразуя их значения. Это позволяет сети моделировать нелинейные зависимости между признаками и улучшает ее способность к аппроксимации сложных функций. Наиболее часто используемая активационная функция в сверточных нейронных сетях — это функция ReLU (rectified linear unit) [31-35]. Она преобразует отрицательные значения в ноль, а положительные значения оставляет без изменений. Математически, функция ReLU определяется следующим образом:

$$\text{ReLU}(x) = \max(0, x). \quad (4)$$

Функция ReLU имеет простую структуру и обладает высокой вычислительной эффективностью, что делает ее популярным выбором для активационной функции в сверточных нейронных сетях (рис. 4б.). Она также помогает в борьбе с проблемой затухания градиента, которая может возникать при обучении глубоких нейронных сетей. Помимо функции ReLU, в сверточных нейронных сетях могут использоваться и другие активационные функции, такие как сигмоидная функция (sigmoid) [36], функция гиперболического тангенса (tanh) [37] и функция softmax [38]. Сигмоидная функция преобразует значения в интервале от 0 до 1, что может использоваться для прогнозирования вероятностей или для бинарной классификации. Функция гиперболического тангенса преобразует значения в интервале от -1 до 1, и она может использоваться для обработки данных с отрицательными значениями. Функция softmax широко используется в последнем слое сверточных нейронных сетей для многоклассовой классификации. Она преобразует выходы сети в вероятности для каждого клас-

са, сумма которых равна 1. Выбор активационной функции зависит от конкретной задачи и требований модели.

Также в CNN может добавляться слой пулинга (pooling layer, PL), сохраняя наиболее значимые признаки и устойчивость к небольшим изменениям входных данных. PL выполняет операцию субдискретизации, сокращая пространственные размеры активаций и уменьшая количество параметров в модели (рис. 4в). Наиболее распространенной операцией пулинга является операция подвыборки (subsampling) [39-40] или операция максимального объединения (max pooling) [41-42]. Операция максимального объединения выполняется путем деления входных данных на неперекрывающиеся окна и выбора максимального значения в каждом окне. Размер окна и шаг перемещения определяются заранее и влияют на степень уменьшения размерности.

Процесс работы слоя пулинга следующий [27, 31]:

- 1) входные данные разделяются на неперекрывающиеся окна. Обычно окна имеют размеры 2 на 2 или 3 на 3;
- 2) в каждом окне выбирается наибольшее значение (для операции максимального объединения) или выполняется другая операция агрегации, такая как усреднение;
- 3) полученные значения формируют новый слой пулинга, в котором каждое значение представляет собой сжатую информацию из соответствующего окна.

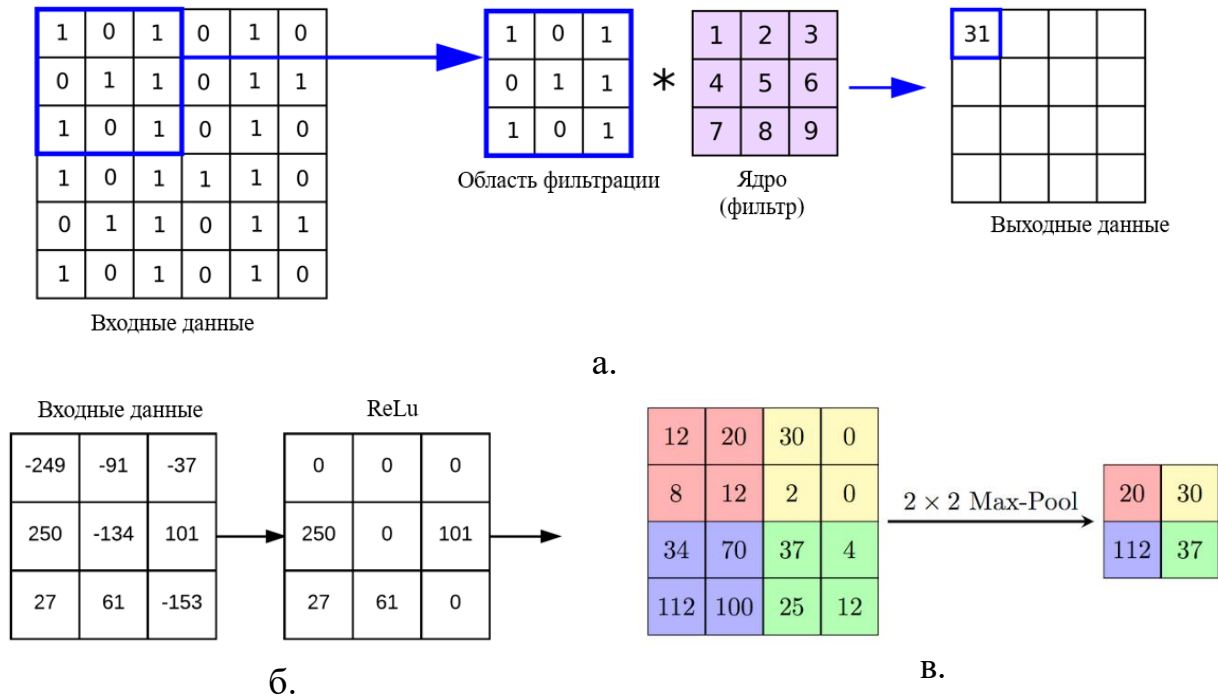


Рис. 4. Пример работы слоев CNN: сверточный слой (а), функция активации (ReLU) (б), пулинг слой (в)

Преимуществами слоя пулинга являются уменьшение размерности, инвариантность к масштабированию и небольшим изменениям, улучшение вычис-

лительной эффективности. Однако, следует отметить, что в некоторых случаях пулинг может приводить к потере информации о точном местоположении объектов в изображении, что может быть нежелательным для некоторых задач. Поэтому некоторые модификации CNN могут использовать слои пулинга с перекрытием.

После прохождения описанных выше слоев данные подаются на полносвязный слой нейронной сети с функцией активации и выполняется классификация данных. После выполнения всех описанных ранее операций данный постулают на каналный кодер, однако данная архитектура может быть построена с применением концепции JSCC, объединяющая источниковое и каналное кодирование для улучшения передачи данных. Это общий порядок действий CNN для обработки изображений, который может быть модифицирован и расширен в зависимости от конкретной архитектуры сети и задачи, но эти шаги представляют основу работы CNN в контексте обработки изображений.

В [43] приведены графики показывающие, что классический показатель помехоустойчивости для инфокоммуникационных систем, а именно битовое отношение сигнал-шум (bit error rate, BER), не является точной оценкой систем с использованием семантики. Основной причиной является возможность оптимизации принятого изображения относительно переданного с целью улучшения его восприятия получателем. Данный подход приводит к формальному различию переданной и принятой информации на бинарном уровне. В следствии чего необходимо использовать другие показатели эффективности для оценки качества восстановленного изображения.

#### 4.1. Показатели эффективности для оценки семантической системы

Одним из основных показателей эффективности для оценки изображений является PSNR (peak signal-to-noise ratio) [42-44]. Данный показатель используется для измерения качества восстановленного сигнала или изображения с учетом искажений или потерь данных. PSNR выражается в децибелах и вычисляется как логарифм отношения максимальной возможной мощности сигнала к среднеквадратической ошибке (mean squared error, MSE) между исходным и восстановленным сигналами. Чем выше значение PSNR, тем ближе восстановленный сигнал к исходному и, следовательно, считается лучшим качеством.

$$PSNR = 10\log_{10}(MAX^2/MSE), \quad (5)$$

где:  $MAX$  – максимальное значение пикселя в зависимости от разрядности. Например, для разрядности 8 бит максимальное значение пикселя составляет 255;  $MSE$  – среднеквадратическая ошибка между исходным и восстановленным сигналами.

Также часто используется показатель SSIM (structural similarity index) [45], который предназначен для измерения структурного сходства между оригинальным и обработанным изображениями, учитывая восприятие человека. SSIM учитывает три основных аспекта визуального восприятия: яркость, контрастность и структуру. Он сравнивает структурные элементы изображений, такие как текстуры, края и градации цвета, и вычисляет сходство между ними.

Высокое значение SSIM указывает на высокое сходство между изображениями, что означает, что они визуальны похожи. SSIM является одним из наиболее распространенных показателей, используемых для оценки качества изображений и видео. Ее применение помогает оптимизировать алгоритмы обработки, сжатия и восстановления изображений, чтобы получить наилучшие результаты с точки зрения визуального сходства с оригиналом.

Помимо вышеприведенных показателей также используются такие показатели как: среднеквадратичная ошибка и средняя абсолютная ошибка (mean absolute error, MAE), которые измеряют среднее или абсолютное отклонение между значениями пикселей исходного и восстановленного изображения. MAE является альтернативным показателем эффективности к MSE, который устойчив к выбросам и может давать представление о средней абсолютной разнице между изображениями [46]. Показатели Precision, Recall и F1-Score, используемые в задачах классификации для оценки производительности модели по определению классов изображений. Precision (точность) измеряет долю правильно классифицированных положительных примеров относительно всех примеров, отмеченных как положительные. Recall (полнота) оценивает долю правильно классифицированных положительных примеров относительно всех истинно положительных примеров. F1-Score является гармоническим средним между Precision и Recall и представляет собой общую оценку производительности классификатора [47].

Данные показатели эффективности могут быть применены для различных задач, таких как восстановление изображений, классификация, детектирование объектов и другие, и помогают в оценке и сравнении работы сверточных нейронных сетей.

## Заключение

В работе проведен анализ исследований, посвященных семантическим инфокоммуникациям. Рассмотрены основополагающие принципы начиная с момента первого упоминания, представлена общая структурная схема семантической системы, представлены типовые подходы реализации нейронных сетей для трех видов передаваемых данных: аудио, текст и графика. Случай передачи графических данных рассмотрен более углубленно с детальным представлением наиболее популяризированного подхода с применением сверточной нейронной сети. Приведены основные показатели оценки качества, описанные в известной литературе.

Введение семантической обработки информации в структуру систем связи обладает высоким потенциалом в части достижения требований, предъявляемым к системам связи будущих поколений, которые стремятся обеспечить повышение достоверности, полноты и контекста информационных сообщений. Семантическая обработка информации также соответствует требованиям адаптивности и гибкости в различных контекстах и областях применения, характерных гетерогенным сетям.



Дальнейшим этапом работы планируется более глубокая проработка математического аппарата ключевых элементов семантической инфокоммуникационной системы. Реализация имитационной модели семантической системы связи для проведения исследований на ее основе с целью повышения достоверности передаваемой информации в семантической инфокоммуникационной системе.

*Исследование выполнено за счет гранта Российского научного фонда № 23-69-10084, <https://rscf.ru/project/23-69-10084/>*

### Литература

1. Morris C. W. Foundations of the theory of signs // The University of Chicago Press. 1938. vol. 1. № 2. P. 59.
2. Shannon E. A mathematical theory of communication // The Bell System Technical Journal. 1948. vol. 27. № 3. P. 379–423.
3. Weaver W., Shannon C. Recent contributions to the mathematical theory of communication. – University of Illinois Press, 1949. – 12 p.
4. Carnap R., Bar-Hillel Y. An Outline of a Theory of Semantic Information. – Massachusetts Institute of Technology: Research Laboratory of Electronics, 1952. – 54 p.
5. Leonard K. Communication Nets: Stochastic Message Flow and Design. – McGraw-Hill, 1964. – 220 p.
6. Шрейдер Ю А. Об одной модели семантической теории информации // Проблемы кибернетики. 1965. С. 233–240.
7. Wilpon J. G., Rabiner L. R., Martin T. An improved word-detection algorithm for telephone-quality speech incorporating both syntactic and semantic constraints // AT&T Bell Laboratories Technical Journal March. 1984. vol. 63. № 3. P. 479–498.
8. Palatucci M., Pomerleau D., Hinton G., Tom Mitchell M. Zero-Shot Learning with Semantic Output Codes // NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems. Vancouver. Canada. 2009. P. 1410–1418. doi: 10.1184/R1/6476456.V1.
9. ГОСТ Р 55062-2012. Информационные технологии (ИТ). Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения. – М.: СтандартИнформ, 2014. – 12 с.
10. Макаренко С. И., Соловьева О. С. Основные положения концепции семантической интероперабельности сетцентрических систем // Журнал радиоэлектроники. 2021. № 4. doi: 10.30898/1684-1719.2021.4.10.
11. Макаренко С. И., Соловьева О. С. Семантическая интероперабельность взаимодействия в сетцентрических системах // Журнал радиоэлектроники. 2021. № 6. doi: 10.30898/1684-1719.2021.6.3.
12. Bao J., Basu P., Dean M., Partridge C., Swami A., Leland W., Hendler J. A. Towards a theory of semantic communication // IEEE Network Science Workshop. New York. USA. 2011. P. 110–117. doi: 10.1109/NSW.2011.6004632.
13. Floridi L. Outline of a theory of strongly semantic information // Minds and machines. 2004. vol. 14. № 2. P. 197221. doi: 10.1023/B:MIND.0000021684.50925.c9.

14. Qin Z., Tao X., Lu J., Li G. Y. Semantic communications: Principles and challenges // arXiv. 2022. vol. 5. P. 1–32. doi: 10.48550/arXiv.2201.01389.
15. Shi G., Gao D., Song X., Chai J., Yang M., Xie X., Li L., Li X. A new communication paradigm: from bit accuracy to semantic fidelity // arXiv. 2021. P. 1–8. doi: 10.48550/arXiv.2101.12649.
16. D’Alfonso S. On quantifying semantic information // Information. 2011. vol. 2. № 1. P. 61–101. doi: 10.3390/info2010061.
17. Kolchinsky A., Wolpert D. H. Semantic information, autonomous agency and non-equilibrium statistical physics // Interface focus. 2018. vol. 8. № 6. P. 17. doi: 10.1098/rsfs.2018.0041.
18. Tong H., Yang Z., Wang S., Hu Y., Semiari O., Saad W., Yin C. Federated learning for audio semantic communication // Frontiers in Communications and Networks. 2021. vol. 2. P. 1–6. doi: 10.3389/frcmn.2021.734402.
19. Weng Z., Qin Z. Semantic communication systems for speech transmission // IEEE Journal on Selected Areas in Communications. 2021. vol. 39. № 8. P. 2434–2444. doi: 10.48550/arXiv.2102.12605.
20. Farsad N., Rao M., Goldsmith A. Deep learning for joint source channel coding of text // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary. Canada. 2018. P. 2326–2330.
21. Wu Y., Schuster M., Chen Z., Le Q. V., Norouzi M., Macherey W., Krikun M., Cao Y., Gao Q., Macherey K. Google’s neural machine translation system: Bridging the gap between human and machine translation // arXiv. 2016. P. 15. doi: 10.48550/arXiv.1609.08144.
22. Xie H., Qin Z., Li G. Y., Juang B. H. Deep learning enabled semantic communication systems // IEEE Transactions on Signal Processing. 2021. vol. 69. P. 2663–2675. doi: 10.1109/TSP.2021.3071210.
23. Lee C H., Lin J W., Chen P H., Chang Y C. Deep learning constructed joint transmission-recognition for internet of things // IEEE Access. 2019. vol. 7. P. 76547–76561. doi: 10.1109/ACCESS.2019.2920929.
24. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas. USA. 2016. P. 770–778. doi: 10.1109/CVPR.2016.90.
25. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate // ICLR. San Diego. USA. 2015. P. 1–15. doi: 10.48550/arXiv.1409.0473.
26. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Polosukhin I. Attention Is All You Need // NIPS. Los Angeles. USA. 2017. P. 1–15. doi: 10.48550/arXiv.1706.03762.
27. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. – М.: ДМК Пресс, 2018. – 652 с.
28. Tolias G., Sicre R., Jégou H. Particular Object Retrieval With Integral Max-Pooling of CNN Activations // International Conference on Learning Representations. San Juan. Puerto Rico. 2016. P. 1–12. doi: 10.48550/arXiv.1511.05879.
29. Khanday O. M., Advandipour S., Lone M. A. Effect of filter sizes on image classification in CNN: A case study on CFIR10 and fashion-MNIST datasets //

International Journal of Artificial Intelligence (IJ-AI). 2021. vol. 10. № 4. P. 872–878. doi: 10.11591/ijai.v10.i4.pp872-878.

30. Dhall A., Ramana M. O., Goecke R., Joshi J., Gedeon T. Video and image based emotion recognition challenges in the wild: Emotiw 2015 // ACM International Conference on Multimodal Interaction. New York. USA. 2015. P. 423–426. doi: 10.1145/2818346.2829994.

31. Паттанаяк С. Глубокое обучение и TensorFlow для профессионалов. Математический подход к построению систем искусственного интеллекта на Python. М.: Диалектика-Вильямс, 2019. 480 с.

32. Agarap A. F. Deep learning using rectified linear units (ReLU) // arXiv. 2018. P. 1–7. doi: 10.48550/arXiv.1803.08375.

33. Karnewar A., Ritschel T., Wang O., Mitra N. ReLU fields: The little non-linearity that could // ACM SIGGRAPH 2022 Conference Proceedings. New York. USA. 2022. P. 1–9. doi: 10.1145/3528233.3530707.

34. Yarotsky D. Error bounds for approximations with deep ReLU networks // Neural Networks. 2017. Vol. 94. P 103–114. doi: 10.1016/j.neunet.2017.07.002.

35. Chen G., Zhang X., Wang Q., Dai F., Gong Y., Zhu K. Symmetrical dense-shortcut deep fully convolutional networks for semantic segmentation of very-high-resolution remote sensing images // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2018. vol. 11. № 5. P. 1633–1644. doi: 10.1109/JSTARS.2018.2810320.

36. Wang Y., Li Y., Song Y., Rong X. The influence of the activation function in a convolution neural network model of facial expression recognition // Applied Science. 2020. vol. 10. № 5. P. 1897. doi: 10.3390/app10051897.

37. Roy S. K., Manna S., Dubey S. R., Chaudhuri B. B. LiSHT: Non-parametric linearly scaled hyperbolic tangent activation function for neural networks // International Conference on Computer Vision and Image Processing. Jaipur. India. 2019. P. 1–15. doi: 10.48550/arXiv.1901.05894.

38. Wydmuch M., Jasinska K., Kuznetsov M., Busa-Fekete R., Dembczyński K. A no-regret generalization of hierarchical softmax to extreme multi-label classification // In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). Red Hook. USA. 2018. P. 6358–6368. doi: 10.48550/arXiv.1810.11671.

39. Liu P., Zhang H., Zhang K., Lin L., Zuo W. Multi-level wavelet-cnn for image restoration // IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Salt Lake City. USA. 2018. P. 773–782. doi: 10.1109/CVPRW.2018.00121.

40. Xi J., Ersoy O. K., Fang J., Cong M., Wu T., Zhao C., Li Z. Wide Sliding Window and Subsampling Network for Hyperspectral Image Classification // Remote Sens. 2021. vol. 13. P. 1290. doi: 10.3390/rs13071290.

41. Rossetti S., Zappia D., Sanzari M., Schaerf M., Pirri F. Max pooling with vision transformers reconciles class and shape in weakly supervised semantic segmentation // ECCV 2022: Computer Vision. Tel Aviv. Israel. 2022. P. 446–463. doi: 10.48550/arXiv.2210.17400.

42. Nurjannah A.F., Kurniasari A. S. D., Sari Z., Azhar Y. Pneumonia Image Classification Using CNN with Max Pooling and Average Pooling // Jurnal RESTI

(Rekayasa Sistem Dan Teknologi Informasi). 2022. vol. 6. № 2. P. 330–338. doi: 10.29207/resti.v6i2.4001.

43. Van Chien T., Phong L. H., Phuc D. X., Hoa N. T. Image Restoration under Semantic Communications // International Conference on Advanced Technologies for Communications (ATC). Ha Noi. Vietnam. 2022. P. 332–337. doi: 10.1109/ATC55345.2022.9943000.

44. Bondzulich B. P., Pavlovic B. Z., Petrovic V. S., Andric M. S. Performance of Peak Signal-to-Noise Ratio Quality Assessment in Video Streaming with Packet Losses // Electronics Letters. 2016. vol. 52. № 6. P. 454–456. doi: 10.1049/el.2015.3784.

45. Horé A., Ziou D. Image quality metrics: PSNR vs. SSIM // 20th International Conference on Pattern Recognition. Istanbul. Turkey. 2010. P. 2366–2369. doi: 10.1109/ICPR.2010.579.

46. Chicco D., Warrens M. J., Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation // Computer science. 2021. vol. 7. № e623. P. 1–24. doi: 10.7717/peerj-cs.623.

47. Stevens E., Antiga L., Viehmann T. Deep Learning with PyTorch. – Manning Publications Co, 2020. – 520 p.

### References

1. Morris C. W. Foundations of the theory of signs. *The University of Chicago Press*, 1938, vol. 1, no. 2, 59 p.

2. Shannon E. A mathematical theory of communication. *The Bell System Technical Journal*, 1948, vol. 27, no. 3, pp. 379–423.

3. Weaver W., Shannon C. *Recent contributions to the mathematical theory of communication*. University of Illinois Press, 1949, 12 p.

4. Carnap R., Bar-Hillel Y. *An Outline of a Theory of Semantic Information*. Massachusetts Institute of Technology: Research Laboratory of Electronics, 1952, 54 p.

5. Leonard K. *Communication Nets: Stochastic Message Flow and Design*. McGraw-Hill, 1964, 220 p.

6. Shrejder Y. A. Ob odnoj modeli semanticheskoy teorii informacii [On one model of semantic information theory]. *Problemy kibernetiki*, 1965, pp. 233–240 (in Russian).

7. Wilpon J. G., Rabiner L. R., Martin T. An improved word-detection algorithm for telephone-quality speech incorporating both syntactic and semantic constraints. *AT&T Bell Laboratories Technical Journal*, 1984, vol. 63, no. 3, pp. 479–498.

8. Palatucci M., Pomerleau D., Hinton G., Tom Mitchell M. Zero-Shot Learning with Semantic Output Codes. *NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems*. Vancouver, Canada, 2009, pp. 1410–1418. doi: 10.1184/R1/6476456.V1.

9. State Standart R 55062-2012. Information Technology (IT). Industrial automation systems and their integration. Interoperability. The main provisions. Moscow, Standartinform, 2014. 12 p.



10. Makarenko S. I., Solovieva O. S. Basic provisions of the concept of semantic interoperability of net-centric systems. *Zhurnal Radioelektroniki* [Journal of Radio Electronics], 2021, no. 4. doi: 10.30898/1684-1719.2021.4.10 (in Russian).
11. Makarenko S. I., Solovieva O. S. Semantic interoperability of interaction of elements in network-centric systems. *Zhurnal Radioelektroniki* [Journal of Radio Electronics], 2021, no. 6. doi: 10.30898/1684-1719.2021.6.3 (in Russian).
12. Bao J., Basu P., Dean M., Partridge C., Swami A., Leland W., Hendler J. A. Towards a theory of semantic communication. *IEEE Network Science Workshop*. New York, USA, 2011, pp. 110–117. doi: 10.1109/NSW.2011.6004632.
13. Floridi L. Outline of a theory of strongly semantic information. *Minds and machines*, 2004, vol. 14, no. 2, pp. 197–221. doi: 10.1023/B:MIND.0000021684.50925.c9.
14. Qin Z., Tao X., Lu J., Li G. Y. Semantic communications: Principles and challenges. *arXiv*, 2022, vol. 5, pp. 1–32. doi: 10.48550/arXiv.2201.01389.
15. Shi G., Gao D., Song X., Chai J., Yang M., Xie X., Li L., Li X. A new communication paradigm: from bit accuracy to semantic fidelity. *arXiv*, 2021, pp. 1–8. doi: 10.48550/arXiv.2101.12649.
16. D’Alfonso S. On quantifying semantic information. *Information*, 2011, vol. 2, no. 1, pp. 61–101. doi: 10.3390/info2010061.
17. Kolchinsky A., Wolpert D. H. Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface focus*, 2018, vol. 8, no. 6, pp. 1–17. doi: 10.1098/rsfs.2018.0041.
18. Tong H., Yang Z., Wang S., Hu Y., Semiari O., Saad W., Yin C. Federated learning for audio semantic communication. *Frontiers in Communications and Networks*, 2021, vol. 2, pp. 1–6. doi: 10.3389/frcmn.2021.734402.
19. Weng Z., Qin Z. Semantic communication systems for speech transmission. *IEEE Journal on Selected Areas in Communications*, 2021, vol. 39, no. 8, pp. 2434–2444. doi: 10.48550/arXiv.2102.12605.
20. Farsad N., Rao M., Goldsmith A. Deep learning for joint source channel coding of text. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Calgary, Canada, 2018, pp. 2326–2330.
21. Wu Y., Schuster M., Chen Z., Le Q. V., Norouzi M., Macherey W., Krikun M., Cao Y., Gao Q., Macherey K. Google’s neural machine translation system: Bridging the gap between human and machine translation. *arXiv*, 2016, 15 p., doi: 10.48550/arXiv.1609.08144.
22. Xie H., Qin Z., Li G. Y., Juang B. H. Deep learning enabled semantic communication systems. *IEEE Transactions on Signal Processing*, 2021, vol. 69, pp. 2663–2675. doi: 10.1109/TSP.2021.3071210.
23. Lee C H., Lin J W., Chen P H., Chang Y C. Deep learning constructed joint transmission-recognition for internet of things. *IEEE Access*, 2019, vol. 7, pp. 76547–76561. doi: 10.1109/ACCESS.2019.2920929.
24. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
25. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. *ICLR*. San Diego, USA, 2015, pp. 1–15. doi: 10.48550/arXiv.1409.0473.



26. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Polosukhin I. Attention Is All You Need. *NIPS*. Los Angeles, USA, 2017, pp. 1–15. doi: 10.48550/arXiv.1706.03762.
27. Goodfellow I., Bengio Y., Courville A. *Deep Learning (Adaptive Computation and Machine Learning series)*. The MIT Press, 2016, 800 p.
28. Tolias G., Sivic R., Jegou H. Particular Object Retrieval With Integral Max-Pooling of CNN Activations. *International Conference on Learning Representations*. San Juan, Puerto Rico, 2016, pp. 1–12. doi: 10.48550/arXiv.1511.05879.
29. Khanday O. M., Dadvandipour S., Lone M. A. Effect of filter sizes on image classification in CNN: A case study on CFIR10 and fashion-MNIST datasets. *International Journal of Artificial Intelligence (IJ-AI)*, 2021, vol. 10, no. 4, pp. 872–878. doi: 10.11591/ijai.v10.i4.pp872-878.
30. Dhall A., Ramana M. O., Goecke R., Joshi J., Gedeon T. Video and image based emotion recognition challenges in the wild: Emotiw 2015. *ACM International Conference on Multimodal Interaction*. New York, USA, 2015, pp. 423–426. doi: 10.1145/2818346.2829994.
31. Pattanayak S. *Pro Deep Learning with TensorFlow. A Mathematical Approach to Advanced Artificial Intelligence in Python*. Apress, 2019, 419 p.
32. Agarap A. F. Deep learning using rectified linear units (ReLU). *arXiv*, 2018, pp. 1–7. doi: 10.48550/arXiv.1803.08375.
33. Karnewar A., Ritschel T., Wang O., Mitra N. ReLU fields: The little non-linearity that could. *ACM SIGGRAPH 2022 Conference Proceedings*. New York, USA, 2022, pp. 1–9. doi: 10.1145/3528233.3530707.
34. Yarotsky D. Error bounds for approximations with deep ReLU networks. *Neural Networks*, 2017, vol. 94, pp. 103–114. doi: 10.1016/j.neunet.2017.07.002.
35. Chen G., Zhang X., Wang Q., Dai F., Gong Y., Zhu K. Symmetrical dense-shortcut deep fully convolutional networks for semantic segmentation of very-high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, vol. 11, no. 5, pp. 1633–1644. doi: 10.1109/JSTARS.2018.2810320.
36. Wang Y., Li Y., Song Y., Rong X. The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Science*, 2020, vol. 10, no. 5, 1897 p., doi: 10.3390/app10051897.
37. Roy S. K., Manna S., Dubey S. R., Chaudhuri B. B. LiSHT: Non-parametric linearly scaled hyperbolic tangent activation function for neural networks. *International Conference on Computer Vision and Image Processing*. Jaipur, India, 2019, pp. 1–15. doi: 10.48550/arXiv.1901.05894.
38. Wydmuch M., Jasinska K., Kuznetsov M., Busa-Fekete R., Dembczyński K. A no-regret generalization of hierarchical softmax to extreme multi-label classification. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*. Red Hook, USA, 2018, pp. 6358–6368. doi: 10.48550/arXiv.1810.11671.
39. Liu P., Zhang H., Zhang K., Lin L., Zuo W. Multi-level wavelet-cnn for image restoration. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Salt Lake City, USA, 2018, pp. 773–782. doi: 10.1109/CVPRW.2018.00121.

40. Xi J., Ersoy O. K., Fang J., Cong M., Wu T., Zhao C., Li Z. Wide Sliding Window and Subsampling Network for Hyperspectral Image Classification. *Remote Sens*, 2021, vol. 13, 1290 p., doi: 10.3390/rs13071290.
41. Rossetti S., Zappia D., Sanzari M., Schaerf M., Pirri F. Max pooling with vision transformers reconciles class and shape in weakly supervised semantic segmentation. *ECCV 2022: Computer Vision*. Tel Aviv, Israel, 2022, pp. 446–463. doi: 10.48550/arXiv.2210.17400.
42. Nurjannah A.F., Kurniasari A. S. D., Sari Z., Azhar Y. Pneumonia Image Classification Using CNN with Max Pooling and Average Pooling. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 2022, vol. 6, no. 2, pp. 330–338. doi: 10.29207/resti.v6i2.4001.
43. Van Chien T., Phong L. H., Phuc D. X., Hoa N. T. Image Restoration under Semantic Communications. *International Conference on Advanced Technologies for Communications (ATC)*. Ha Noi, Vietnam, 2022, pp. 332–337. doi: 10.1109/ATC55345.2022.9943000.
44. Bondzulich B. P., Pavlovic B. Z., Petrovic V. S., Andric M. S. Performance of Peak Signal-to-Noise Ratio Quality Assessment in Video Streaming with Packet Losses. *Electronics Letters*, 2016, vol. 52, no. 6, pp. 454–456. doi: 10.1049/el.2015.3784.
45. Horé A., Ziou D. Image quality metrics: PSNR vs. SSIM. *20th International Conference on Pattern Recognition*. Istanbul, Turkey, 2010, pp. 2366–2369. doi: 10.1109/ICPR.2010.579.
46. Chicco D., Warrens M. J., Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Computer science*, 2021, vol. 7, no. 623, pp. 1–24. doi: 10.7717/peerj-cs.623.
47. Stevens E., Antiga L., Viehmann T. *Deep Learning with PyTorch*. Manning Publications Co, 2020. 520 p.

Статья поступила 15 июля 2023 г.

### Информация об авторах

*Вершков Павел Андреевич* – аспирант кафедры инфокоммуникаций. Московский авиационный институт (национальный исследовательский университет). Область научных интересов: инфокоммуникационные технологии и системы связи. E-mail: enemy12340@gmail.com

*Казачков Виталий Олегович* – кандидат технических наук. Доцент кафедры инфокоммуникаций. Московский авиационный институт (национальный исследовательский университет). Область научных интересов: инфокоммуникационные технологии и системы связи. E-mail: kazachkov.vo@gmail.com

*Шевцов Вячеслав Алексеевич* – доктор технических наук, профессор. Заведующий кафедрой инфокоммуникаций. Московский авиационный институт (национальный исследовательский университет). Область научных интересов: инфокоммуникационные технологии и системы связи. E-mail: vs.gtumai@gmail.com

Адрес: 125993, Россия, г. Москва, Волоколамское шоссе, д. 4.

## Semantic infocommunications as the next stage in the development of communication

P. A. Vershkov, V. O. Kazachkov, V. A. Shevtsov

**Purpose.** Currently, telecommunications systems are experiencing tremendous rates of development. At the same time, one of the reasons is the active introduction of artificial neural networks and deep learning technologies. This feature makes it possible to implement a fundamentally new approach based on the extraction of semantics in the construction of communication systems applicable to various fields from civil to military orientation, focusing on the basic meaning of the transmitted information in contrast to the traditional form of binary stream transmission characteristic of systems of previous generations. The task of the work is to study semantic info communications. **The aim of the work** is to form an analytical review based on the aggregation and structuring of information about semantic infocommunication systems from the origin of the concept and its foundations to modern implementations of various approaches, depending on the type of information transmitted at the application level. **Methods.** The implementation of semantic infocommunications is based on the apparatus of higher mathematics. In particular, linear algebra is the main tool of mathematical calculations necessary for the implementation of an artificial neural network. A neural network is responsible for extracting semantics from transmitted data and is a key element of signal processing in the structure of a semantic infocommunication system. **Results and novelty.** An element of the practical novelty of the work is the systematized features of constructing a communication system with the semantic level of information processing. In particular, a general block diagram is given, the main implementation approaches for various types of data, such as audio, text and graphic data, are considered. The main emphasis is placed on a detailed analysis of the case of transmission and reception of graphic data. The structure and typical approach of semantics extraction based on convolutional neural network for image processing are derived. The main indicators of noise resistance are structured. **Practical relevance.** The presented information will be useful in scientific and engineering practice in the study and development of semantic infocommunications.

**Key words:** semantic infocommunications, semantic coding, heterogeneous communication systems, cellular communication, communication system architecture, neural networks, semantic interaction, convolutional neural networks.

### Information about Authors

*Pavel Andreevich Vershkov* – postgraduate student of the Department of Infocommunications. Moscow Aviation Institute (National Research University). Research interests: infocommunication technologies and communication systems. E-mail: enemy12340@gmail.com

*Vitaly Olegovich Kazachkov* – Ph.D. of Engineering Sciences. Associate Professor of the Department of Infocommunications. Moscow Aviation Institute (National Research University). Research interests: infocommunication technologies and communication systems. E-mail: kazachkov.vo@gmail.com

*Vyacheslav Alekseevich Shevtsov* – Dr. habil. of Engineering Sciences, Full Professor. Head of the Department of Infocommunications. Moscow Aviation Institute (National Research University). Research interests: infocommunication technologies and communication systems. E-mail: vs.gtumai@gmail.com

Address: Russia, 125993, Moscow, Volokolamsk highway, 4.