

**П. А. Степанов, Д. Е. Пальчунов, А. А. Мирзагитов**

*Новосибирский государственный университет  
ул. Пирогова, 2, Новосибирск, 630090, Россия*

*Институт математики им. С. Л. Соболева СО РАН  
пр. Акад. Коптюга, 4, Новосибирск, 630090, Россия*

*palch@math.nsc.ru; stefan.nsk@gmail.com; azat90@gmail.com*

## **МЕТОДЫ АНАЛИЗА ДИАЛОГОВ, ОСНОВАННЫЕ НА ТЕОРИИ РЕЧЕВЫХ ДЕЙСТВИЙ\***

Работа посвящена проблеме семантического анализа контента социальных сетей. Решается задача выявления диалогов и полилогов, при помощи которых участники общения занимаются организацией некоторых совместных действий. Цель разработанной программной системы – раннее обнаружение готовящихся противоправных действий и террористических актов, выявление криминальных групп. Предложенные методы основаны на применении теории речевых действий. С помощью языка описания лингвистических шаблонов решается проблема выявления фраз русского языка, содержащих речевые действия побуждения. Показано, что полилоги, содержащие большое количество побуждений, – это в точности те полилоги, при помощи которых в социальных сетях организуются совместные действия.

*Ключевые слова:* теория речевых действий, диалог, полилог, социальные сети, извлечение знаний, анализ текстов естественного языка, лингвистические шаблоны.

### **Введение**

Статья посвящена разработке системы, анализирующей диалоги и полилоги, содержащиеся в социальных сетях, и определяющей, имеет ли место факт обсуждения планов или договоренности собеседников о каких-либо будущих действиях. Эта задача решается с целью выявления готовящихся противоправных действий и террористических актов.

Для решения поставленной задачи была применена теория речевых действий. В процессе разработки сформулирована гипотеза, что искомые полилоги должны содержать высокое количество речевых действий вида «побуждение». Исходя из этой гипотезы был разработан алгоритм поиска требуемых полилогов, основанный на вычислении доли речевых действий побуждений относительно общего количества фраз. Для поиска речевых действий использовался язык описания лингвистических шаблонов, позволяющий описывать требования к грамматическим структурам, а затем производить выявление таких структур в тексте.

Тестовая выборка содержала в основном диалоги при незначительном количестве полилогов. Поэтому далее будем говорить только о диалогах. Источником диалогов была выбрана популярная социальная сеть «ВКонтакте», содержащая более 100 миллионов пользователей. После тестирования системы на извлеченных из социальной сети текстах был получен удов-

---

\* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 14-07-00903\_а, а также при финансовой поддержке Минобрнауки России, задание № 2014/139 на выполнение государственных работ в сфере научной деятельности в рамках базовой части.

*Степанов П. А., Пальчунов Д. Е., Мирзагитов А. А.* Методы анализа диалогов, основанные на теории речевых актов // Вестн. Новосиб. гос. ун-та. Серия: Информационные технологии. 2014. Т. 12, вып. 4. С. 102–111.

летворительный результат. Было найдено 83 % всех диалогов, содержащих договоренности. Количество ошибок в выдаче составило 19 %.

В настоящее время все большее значение приобретают информационные войны, ведущиеся с использованием компьютерных сетей, в том числе в сети Интернет. Информационные войны используют все новые онлайн-сервисы: социальные сети, блоги, электронные журналы и т. д. Такие сервисы не только объединяют людей и дают дополнительные возможности по обмену информацией, свободному ее обсуждению и самовыражению, они также служат для дезинформации, пропаганды, управления мнениями, т. е. для ведения информационной войны.

С развитием информационно-коммуникационных технологий за последние десять лет существенно возросла важность ресурсов нового типа, онлайн-социальных сетей, как средства распространения мнений, влияющих на действия пользователей сети. Социальные сети, как часть Интернета, в первую очередь являются источниками информации. Это связано с тем, что социальные сети, такие как Facebook, VKontakte, способствуют не только организации социальных коммуникаций между людьми, но и реализации их базовых социальных потребностей. Привлекательность социальных сетей обуславливается возможностью выразить свое мнение, узнать мнение друзей, повлиять на их мнение.

Социальные сети стали существенным инструментом информационного влияния, в том числе в целях манипулирования личностью, социальными группами и обществом в целом. Именно поэтому в последние годы социальные сети становятся более популярными в информационной войне, чем телевидение, газеты и другие СМИ.

Сетевые средства используются в том числе и для проведения психологических операций [1]. С началом любого локального конфликта резко повышается активность пользователей социальных сетей – сторонников из обеих противостоящих сторон в Интернете. Цель организаторов сообществ в социальных сетях заключается в поддержке и консолидации в сети тех пользователей, которые поддерживают их политические и стратегические взгляды. Главная задача сообщества – координация разъяснительной работы по своей позиции, при этом используются классические методы информационной пропаганды, в том числе дезинформация.

Поскольку социальные сети становятся мощным средством распространения информации, возникают вопросы контроля над циркулирующей информацией. Это связано с тем, что определенные информационные воздействия, которые распространяются через социальные сети, существенно меняют оценки и взгляды пользователей. Таким образом, социальная сеть потенциально является ареной информационного противоборства [2].

### **Анализ текстов естественного языка**

*Язык описания лингвистических шаблонов.* Для обеспечения возможности поиска в текстах естественного языка грамматических структур, организованных определенным способом, был разработан язык описания лингвистических шаблонов [3]. Этот язык был опробован ранее для решения различных задач анализа текста [4–6].

Язык описания лингвистических шаблонов задает требования к искомым участкам текста. Минимальным элементом, который может быть описан средствами данного языка, является слово. Конструкции, описывающие требования к слову, выступают базовыми элементами, из которых затем строятся условия на более сложные грамматические структуры. Условие на слово выглядит следующим образом:

*<«быть»: "ГЛАГ": "пвл. ">*

В приведенной выше записи требования, описывающие слово, представляют собой кортеж из одной или более строк. Первая строка определяет условие на исходную форму слова. Остальные строки определяют набор грамматических характеристик, которые должны быть реализованы искомым словом. В кортеже эти характеристики записываются при помощи заранее определенных сокращений. В приведенном выше примере описывается глагол «быть» в повелительном наклонении.

Вторым базовым объектом, описываемым разработанным языком, является лингвистический шаблон. Шаблон – это параметризованный набор условий на искомую структуру текста. Простейшим шаблоном является кортеж, содержащий условия на слово. Особенностью шаблона является то, что он может быть параметризован. Объявленные в шаблоне параметры могут быть в дальнейшем использованы для построения более сложных структур, в которых предъявляются условия к параметрам шаблона. Например, рассмотрим следующий шаблон:

$$\text{NOUN}(\text{род}, \text{число}, \text{падеж}) := \\ \langle *:\text{"СУЩ"}:\$род:\$число:\$падеж \rangle$$

Этот шаблон описывает существительное, он параметризован значениями рода, числа и падежа. В дальнейшем данный шаблон может быть использован для построения более сложных шаблонов, в которых условия предъявляются на значения рода, числа или падежа существительного. Например:

$$\text{RELATION}(\text{род}, \text{число}, \text{падеж}) := \\ \text{NOUN}(\$род, \$число, \$падеж) \text{NOUN}(*, *, \langle \text{рд.п.} \rangle)$$

В приведенном примере описывается связь подчинения, например, как в словосочетании «директор фабрики». Шаблон RELATION определяется как два идущих подряд шаблона NOUN. Первый шаблон имеет значения рода, падежа и числа равные заданным параметрам, а второй может иметь произвольные значения рода и числа, но должен иметь значение падежа равное «рд. п.», т. е. «родительный падеж». Определенный таким образом шаблон RELATION будет описывать конструкции типа «параметр шаблона», «платье невесты», «свет звезд», и т. п.

Язык лингвистических шаблонов унаследовал набор средств для конструирования лингвистических шаблонов от широко распространенного языка регулярных выражений.

*Конструкция «Последовательность».* Любые два выражения языка шаблонов могут быть записаны последовательно. Например, как в приведенном ранее примере шаблона RELATION:

$$\text{NOUN}(\$род, \$число, \$падеж) \text{NOUN}(*, *, \langle \text{рд.п.} \rangle)$$

Полученное выражение будет описывать конструкции, состоящие из двух последовательно идущих участков текста, первый из которых описывается первым выражением, а второй – вторым.

*Конструкция «Альтернатива».* Любое множество выражений может быть сформировано в конструкцию альтернатив. Например:

$$[ \langle *:\text{"СУЩ"}:\rangle\text{им.п.}\rangle \mid \langle *:\text{"СУЩ"}:\rangle\text{вн.п.}\rangle ]$$

Приведенная в примере конструкция описывает существительные, стоящие в именительном или винительном падеже.

*Конструкция «Множественность».* Любое выражение языка шаблонов может быть использовано для конструирования множественности. Например:

$$\langle *:\text{"ПРИЛ"} \rangle *$$

Полученное выражение будет описывать набор подряд идущих имен прилагательных.

*Интерпретатор языка шаблонов.* Для поиска участков текста, подходящих под условия, описанные в некотором лингвистическом шаблоне, был реализован интерпретатор. Интерпретатор является Web-приложением, принимающим на вход:

- 1) входной текст;
- 2) набор шаблонов и их определений;

3) имя шаблона, который необходимо искать.

После подтверждения ввода интерпретатор отображает множество найденных участков текста, соответствующих шаблону.

*Теория речевых действий* восходит к Остину и Серлю [7–10]. Логическая формализация теории речевых действий была проделана Серлем и Вандервекенем в разработанной ими иллокутивной логике [11–14].

Идея рассматривать фразы естественного языка как действия, совершаемые говорящим и направленные на изменение реального мира, принадлежит Остину [7; 9]. Такие действия Остином были названы речевыми действиями. По существу, такой подход к формальному описанию естественного языка основан на юридической парадигме. Говорящий при произнесении фразы естественного языка либо берет на себя ответственность за ее верность или выполнение (обещание), либо накладывает ответственность на слушающего (просьба, приказ) или на саму фразу (сообщения, утверждения). Идея наложения ответственности за успех или удовлетворение речевого действия на говорящего или его собеседника является центральной в классификации речевых действий Остина – Серля.

Речевые действия делятся на следующие классы.

- Репрезентативы. Цель такого высказывания – сообщить о прошлом, текущем или будущем состоянии реального мира. При этом ответственность за истинность произносимого высказывания накладывает на само высказывание: оно должно соответствовать реальному миру. Этот вид речевых действий очень близок к предложениям классической логики предикатов; к ним, в отличие от других речевых действий, применима оценка истинно или ложно. Заметим, что репрезентатив «Завтра будет хорошая погода» принципиально отличается от другого речевого действия: «Я обещаю, что завтра будет хорошая погода». Второе речевое действие является директивой – обещанием.

- Директивы. Цель такого речевого действия – побудить собеседника к определенным действиям или поступкам. Говорящий может приказывать собеседнику, запретить, пригласить, посоветовать, попросить его, спросить (т. е. попросить или приказывать ответить) и т. д. При этом ответственность за выполнение (в данной терминологии, «удовлетворение») речевого действия возлагается на собеседника.

- Комиссивы. Цель речевого действия – взятие на себя говорящим ответственности за определенное будущее состояние мира. Говорящий может пообещать, дать гарантии, дать свое согласие. Ответственность за удовлетворение данного речевого действия возлагается на говорящего.

Также выделяются еще два специфических вида речевых действий, при совершении которых ни на кого никакой ответственности не налагается.

- Экспрессивы. Цель – выражение говорящим своего эмоционального состояния: радости, сожаления, сочувствия, благодарности, стыда (когда говорящий извиняется) и проч.

- Декларативы. Цель декларатива – изменить реальный мир, причем это изменение происходит самим фактом произнесения данной фразы. Например, объявление войны, объявление судьей вердикта, высказывание своей оценки членом жюри, заявление своего согласия на сделку в нотариальной конторе или согласия на брак в ЗАГСе. Для этого класса речевых действий наиболее отчетливо проявляется юридическая парадигма, лежащая в основе всей классификации.

Для последних двух речевых действий не возникает вопрос об ответственности, поскольку они выполняются сразу фактом своего произнесения: как выражение эмоций, так и формальное заявление своего согласия.

Данная классификация речевых действий имеет ряд недостатков. Во-первых, вид речевого действия однозначно и чисто синтаксически определяется по соответствующему иллокутивному глаголу: сообщать, утверждать, уведомлять и др. для репрезентативов; приказывать, просить, спрашивать и др. для директивов; обещать, обязываться, соглашаться и др. для комиссивов; поздравлять, сочувствовать, извиняться и др. для экспрессивов; заявлять, объявлять, назначать и др. для декларативов. Таким образом, как было показано, фразы «Я утверждаю, что завтра будет хорошая погода» и «Я обещаю, что завтра будет хорошая погода» относятся к разным классам речевых действий – к репрезентативам и комиссивам соответственно, хотя, очевидно, и семантически и даже прагматически означают одно и то же.

Как следствие этого, во-вторых, предполагается, что каждая фраза относится ровно к одному классу речевых действий: исходя из образующего ее сказуемого – иллокутивного глагола. Поэтому если председатель жюри объявит участнику соревнований: «Я поздравляю Вас с победой в соревновании!», то это будет экспрессив. Если он скажет: «Я объявляю Вас победителем соревнований. Я поздравляю Вас с победой!» это будет два речевых действия: декларатив и экспрессив. Если же он скажет: «Я сообщаю, что Вы победили в соревновании. Я поздравляю Вас с победой!», то это уже будут речевые действия репрезентатив и экспрессив. При этом очевидно, что семантически три этих выступления председателя жюри друг от друга не отличаются. Также легко видеть, что по существу все три выступления предназначены для достижения двух целей: сообщить, кто победил в соревновании, и поздравить его с победой. Иначе говоря, на самом деле все они являются композицией сообщения (репрезентатива) и выражения эмоций (экспрессива), независимо от грамматической конструкции и наличия одного или двух предложений.

Конечно, классический подход к определению видов речевых действий очень удобен в силу своей крайней простоты: для классификации речевого действия достаточно просто установить тип иллокутивного глагола – сказуемого. Однако чуда не происходит: в силу такой примитивности от нас практически полностью ускользает реальная семантика фраз естественного языка, осуществляющих речевые действия.

Для достижения целей, указанных в начале данной статьи: мониторинг социальных сетей и выявление ситуаций, в которых несколько участников обсуждений договариваются о совершении в дальнейшем некоторых совместных действий (возможно, криминальных) нам необходим не чисто синтаксический, а семантический анализ текстов. Конечно, семантику мы извлекаем при помощи анализа синтаксиса, но нам необходимо в конечном счете определять именно семантические, а не синтаксические инварианты. Поэтому для достижения наших целей мы будем использовать не классический, а усовершенствованный подход к определению целей речевых действий [15–17].

В рамках данного подхода, во-первых, мы отказываемся от отождествления видов речевых действий с видами иллокутивных глаголов. Во-вторых, мы предполагаем, что одно речевое действие может преследовать несколько целей. Мы считаем, что речевое действие состоит из одной или нескольких компонент, имеющих разные цели и поэтому относящихся к разным видам речевых действий. Таким образом, уменьшая количество видов компонент речевых действий, мы получаем достаточно большое количество классов сложных речевых действий, состоящих из нескольких компонент.

И, наконец, мы отходим от амбициозного постулата Остина о том, что слово (речевое действие) может менять мир. Слово действительно может менять реальный мир, но только в крайне ограниченном наборе ситуаций, имеющих чисто юридическую окраску: объявление войны, вынесение приговора и проч. Рассмотрение таких ситуаций малоинтересно с практической точки зрения: для них и без анализа естественного языка и так все понятно, в силу опять же полной юридической формализованности этих ситуаций.

Тем не менее, произнесение слов все-таки кое-что меняет во внешнем мире, а именно: оно меняет собеседника (или собеседников). Других изменений реального мира произнесение фраз, как правило, не производит.

Поэтому за основу классификации видов простейших речевых действий берется модель участников речевого общения.

Мы выделяем три составляющих участника общения. Первая – это знания, система представлений об окружающем мире. Вторая – это намерения, желания, мотивы и цели. Третья – это представление о происходящем прямо сейчас, о контексте диалога, о смысле используемых в настоящий момент времени терминах (используемая онтология), об эмоциональном отношении собеседника. Заметим, что эти три части «ума» участника диалога фактически не пересекаются. Поэтому компонентами цели речевого действия мы можем считать воздействие на каждую из этих частей по отдельности.

Таким образом, мы имеем три вида целей речевых действий. Первая цель – изменение знаний и представлений собеседника, сообщение ему новой информации. Речевые действия (компоненты речевых действий), преследующие эту цель, мы называем сообщениями.

Вторая цель речевых действий – изменение намерений собеседника, побуждение его к определенным действиям или, наоборот, к отказу от определенных действий. Такие речевые действия называются побуждениями.

И, наконец, третья цель речевых действий – изменение представления собеседника о происходящем прямо сейчас, о содержательном и эмоциональном контексте диалога. Речевые действия, направленные на достижение этой цели, мы называем заявлениями. Например, один собеседник может заявить о своем хорошем или плохом отношении к тому, что говорит другой собеседник. Также он может уточнить смысл используемых понятий, заявить о том, что в данном контексте значения используемых им терминов отличаются от общеупотребительных и относятся к определенному жаргону.

Речевое действие может состоять из компонент одного, двух или всех трех видов. Например, приведенная выше фраза председателя жюри при подведении итогов: «Я поздравляю Вас с победой в соревновании!» является композицией двух видов речевых действий: сообщения о победе и заявления о поздравлении – выражения позитивных эмоций в момент объявления результатов. Речевое действие «Ты должен посетить лекцию, которая состоится завтра в таком-то месте» является композицией сообщения о лекции и побуждения ее посетить. Речевое действие «Я рад сообщить, что завтра ты должен прийти на заседание ученого совета для вручения тебе почетной грамоты» является композицией сообщения о вручении почетной грамоты, побуждения к приходу и заявления о позитивных эмоциях говорящего.

Для целей определения диалогов и полилогов, в которых их участники договариваются о совершении определенных действий, нас будут интересовать речевые действия, имеющие компоненту «побуждение». Нам необходимо, во-первых, достаточно точно определять такие речевые действия. Во-вторых, необходимо выяснить, какой процент побуждений является необходимым и достаточным для идентификации таких диалогов и полилогов.

### **Извлечение фактов договоренностей собеседников**

Задачей, в рамках которой была проведена данная работа, был поиск диалогов в социальных сетях, в которых собеседники договариваются о чем-то, строят планы. Согласно описанной выше теории речевых действий достаточно очевидно, что подобные диалоги будут в большом количестве содержать речевые действия побуждения, имеющие своей целью изменить намерения и желания собеседника.

*Построение корпуса диалогов.* В качестве входных данных необходимо было создать размеченный корпус диалогов, так как мы столкнулись с такой проблемой, что в открытом виде на русском языке не существует подобных ресурсов.

На первом этапе работы был набран корпус диалогов из социальной сети «ВКонтакте». Был набран корпус из личных диалогов между двумя собеседниками по 10–20 сообщений. Часть диалогов была получена непосредственно из личной переписки с согласия участников этой переписки.

Также был реализован модуль, осуществляющий извлечение публичных полилогов из текстов общения произвольного количества собеседников. Данный модуль при тестировании извлекал до 100 диалогов и полилогов из текстов. В зависимости от сообщества, количество собеседников, тематика и объем могли быть различны. Для тестирования в большинстве случаев использовались сообщества, содержащие радикальные, экстремистские и националистические материалы и соответствующий состав участников.

Поскольку первоначальной целью было получение диалогов, а не сообщений, то был установлен следующий фильтр: если сообщение содержит более пяти комментариев, то оно выкачивается. Формат «сообщение и несколько комментариев к нему» наиболее близок к формату «диалог», единственная разница – комментарии могут оставлять несколько людей, в то время как диалог ведется двумя собеседниками. Тем не менее, в результате анализа собранных данных было установлено, что формат «сообщение и несколько комментариев к нему» хорошо ложится в рамки структуры обычного диалога. Было обработано порядка десяти групп, в результате было получено порядка тысячи диалогов.

*Критерии поиска.* В процессе анализа корпуса полученных диалогов была выдвинута гипотеза о том, что в искомым диалогах, содержащих договоренности, будет также содержать-

ся много речевых действий типа «побуждение». Эта гипотеза хорошо согласуется с определением речевого действия «побуждение», целью которого является изменение намерений, целей, желаний собеседника.

Данная гипотеза была проверена на практике. Поиск речевых действий а побуждений производился при помощи языка описания лингвистических шаблонов.

Следующие лингвистические шаблоны были использованы для поиска:

```
// просьба, приказ. глагол в повелительном наклонении
TEMPLATE_01() :=
  <*: "ГЛАГ": "вел. ">;
// декларация планов. конструкции в будущем
// времени. будешь делать
TEMPLATE_02() :=
  <"": "ГЛАГ": "будущ.вр." > <*: "ИНФ">;
// декларация планов. конструкции в будущем
// времени, собираешься делать
TEMPLATE_03() :=
  <"собира": "ГЛАГ": "наст.вр." > <*: "ИНФ">;
// декларация планов. конструкции в будущем
// времени. планировал делать
TEMPLATE_04() :=
  <"планир": "ГЛАГ": "прош.вр." > <*: "ИНФ">;
// декларация планов. конструкции в будущем
// времени. собирался делать
TEMPLATE_05() :=
  <"собира": "ГЛАГ": "прош.вр." > <*: "ИНФ">;
// косвенное указание. предлагаю, предложено сделать
TEMPLATE_06() :=
  ([ <"предлага": "ГЛАГ": "наст.вр." > |
  <"предлож": "ПРИЧ": "прош.вр." > ])
  <*: "ИНФ">;
// косвенное указание. советую, советовал сделать
TEMPLATE_07() :=
  ([ <"совет": "ГЛАГ": "наст.вр." > |
  <"совет": "ГЛАГ": "прош.вр." > ])
  <*: "ИНФ">;
// косвенное указание. рекомендую, рекомендовал сделать
TEMPLATE_08() :=
  ([ <"рекоменд": "ГЛАГ": "наст.вр." > |
  <"рекоменд": "ГЛАГ": "прош.вр." > ]) <*: "ИНФ">;
// косвенное указание. следует, нужно, надо сделать
TEMPLATE_09() :=
  ([ <"след": "ГЛАГ": "безл." > |
  <"надо": "ПРЕДИК": "наст.вр." > |
  <"нужно": "ПРЕДИК": "наст.вр." > ]) <*: "ИНФ">;
```

Приведенные лингвистические шаблоны соответствуют следующим критериям поиска.

1. Просьба, приказ. Глагол в повелительном наклонении.
2. Декларация планов. Конструкции «будешь делать», «собираешься делать», «планировал сделать», «собирался сделать».
3. Косвенные указания. Конструкции «предлагаю / предложено сделать», «советую / советовал сделать», «рекомендую / рекомендовал сделать», «следует, нужно, надо сделать».

Произведенный поиск речевых действий среди корпуса диалогов в результате дал удовлетворительные показатели эффективности данного набора лингвистических шаблонов. Пол-

нота поиска речевых действий «побуждение» с помощью шаблонов составила 88 %, доля ошибок среди результатов – 18 %.

Таким образом, была решена задача поиска речевых действий побуждений в текстах русского языка.

*Извлечение фактов договоренностей.* Для определения факта договоренности собеседников в некотором диалоге сначала производился поиск побуждений в тексте диалога. Затем вычислялась доля найденных речевых действий относительно общего количества фраз в диалоге. Если доля побуждений превосходила определенный порог, делался вывод, что диалог содержит факт договоренности собеседников об осуществлении некоторых действий.

Экспериментально установлено, что требуемое значение порога – 10 %. Используя данное значение порога, мы выявили 83 % всех диалогов, содержащих договоренности.

Доля побуждений, %	Количество диалогов	
	содержащих договоренности	не содержащих договоренности
0–10	6	61
10–15	10	5
15–20	3	1
20–25	3	0
25–100	13	1

Как видно из таблицы, при использовании порога в 10 % для определения диалогов с договоренностями в отфильтрованный результат попадает 83 % всех диалогов с договоренностями. При этом в результате содержится 19 % диалогов, которые алгоритмом были найдены ошибочно.

## Заключение

Разработана программная система, осуществляющая автоматическое извлечение публичных диалогов из сообществ социальной сети «ВКонтакте».

Создан набор лингвистических шаблонов, с помощью которых можно осуществлять поиск речевых действий вида «побуждение». Осуществлен поиск побуждений в корпусе диалогов. Подтверждена гипотеза о наличии достаточно большого количества побуждений в диалогах, в которых осуществляется организация совместных действий.

В дальнейшем планируется построить представительный корпус диалогов и полилогов для разработки и тестирования алгоритмов распознавания текстов, содержащих организацию совместных действий. Также предполагается разработка методов анализа диалогов и полилогов с использованием онтологий. Цель такого анализа – определение принадлежности текстов определенной тематике, в частности, террористической, экстремистской, тематике, связанной с совершением или планированием противоправных действий.

## Список литературы

1. Губанов Д. А., Новиков Д. А., Чхартушвили А. Г. Социальные сети: модели информационного влияния, управления и противоборства. М.: Физматлит, 2010.
2. Трахтенгерц Э. А. Компьютерные технологии информационного управления в конфликтных ситуациях // Информационные технологии. Приложение. 2009. № 1. С. 1–32.
3. Степанов П. А. Язык описания лингвистических шаблонов // Материалы Всерос. конф. с междунар. участием «Знания – Онтологии – Теории». 2013. Т. 2. С. 136–145.
4. Власов Д. Ю., Пальчунов Д. Е., Степанов П. А. Автоматизация извлечения отношений между понятиями из текстов естественного языка // Вестн. Новосиб. гос. ун-та. Серия: Информационные технологии. 2010. Том 8, вып. 3. С. 23–33.
5. Пальчунов Д. Е., Степанов П. А. Применение теоретико-модельных методов извлечения онтологических знаний в предметной области информационной безопасности // Программная инженерия. 2013. № 11. С. 8–16.



6. *Степанов П. А.* Автоматизация обработки текстов естественного языка. // Вестн. Новосибир. гос. ун-та. Серия: Информационные технологии. 2013. Т. 11, вып. 2. С. 106–112.
7. *Austin J. L.* How to do things with words // Set of lectures. Oxford: Oxford Univ. Press, 1962.
8. *Searle J. R.* Speech acts: an essay in the philosophy of language. Cambridge: Cambridge Univ. Press, 1969.
9. *Остин Дж.* Слово как действие // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 22–129.
10. *Серль Дж. Р.* Классификация иллокутивных актов // Новое в зарубежной лингвистике. М., 1986. Вып. 17: Теория речевых актов. С. 170–194.
11. *Searle J. R. Vandervecken D.* Foundations of Illocutionary Logic. Cambridge Univ. Press, 1985.
12. *Серль Дж. Р., Вандервекен Д.* Основные понятия исчисления речевых актов // Новое в зарубежной лингвистике. М., 1986. Вып. 18: Логический анализ естественного языка. С. 242–264.
13. *Vandervecken D.* Meaning and Speech Acts. Cambridge Univ. Press, 1990. Vol. 1: Principles of Language Use.
14. *Vandervecken D.* Meaning and Speech Acts. Cambridge Univ. Press, 1991. Vol. 2: Formal Semantics of Success and Satisfaction.
15. *Пальчунов Д. Е.* О логическом анализе естественного языка // Теория вычислений и языки спецификаций. Новосибирск, 1995. Вып. 152: Вычислительные системы. С. 61–75.
16. *Пальчунов Д. Е.* Алгебраическое описание смысла высказываний естественного языка // Вычислительные системы. Новосибирск, 1997. Вып. 158. С. 127–148.
17. *Pal'chunov D. E.* Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache // GABEK. Verarbeitung und Darstellung von Wissen. J. Zelger, M. Maier (Hrsg.). Innsbruck; Wien: STUDIEN Verlag, 1999. S. 310–326.

*Материал поступил в редколлегию 18.12.2014*

**P. A. Stepanov, D. E. Palchunov, A. A. Mirzagitov**

*Novosibirsk State University  
2 Pirogov Str., Novosibirsk, 630090, Russian Federation*

*Institute of Mathematics SB RAS  
4 Koptyug Ave., Novosibirsk, 630090, Russian Federation*

*palch@math.nsc.ru, stefan.nsk@gmail.com, azat90@gmail.com*

### **METHODS OF ANALYSIS OF DIALOGUES BASED ON THE THEORY OF SPEECH ACTS**

The paper is devoted to the semantic analysis of content of the social networks. The problem of identifying dialogues and polylogues using which participants communicate to organize some joint action is solving. The aim of the developed software system is early detection of imminent unlawful acts and acts of terrorism, detection of criminal groups. The proposed methods are based on the theory of speech acts. The problem of identifying Russian language phrases containing speech acts «directives» is solved using the language of linguistic patterns description. It is shown that polylogues containing a large number of directives are exactly the polylogues using which the joint actions are organized in the social networks.

*Keywords:* theory of speech acts, dialogue, polylogue, social networks, knowledge extraction, natural language processing, linguistic patterns.

## References

1. Gubanov D. A., Novikov D. A., Chartishvili A. G. Social networks: model of informational influence, control and confrontation. Moscow, Physical and mathematical literature publisher, 2010, 228 p.
2. Trahtengerz E. A. Computer Technology of Information Management in Conflict Situations. *Information Technology. Application*, 2009, №1, p. 1–32.
3. Stepanov P. A. Linguistic patterns description language. *All-Russian Conference with international participation «Knowledge Ontology-Theory»*, 2013. Vol. 2, p. 136-145.
4. Vlasov D. J., Pal'chunov D. E., Stepanov p. A. Automating the extraction of relations between concepts from natural language texts. *Vestnik Novosib. gos. uni.* Block: Information technology, 2010. Vol. 8, № 3, p. 23-33.
5. Pal'chunov D. E., Stepanov p. A. The use of model-theoretic methods of extraction of ontological knowledge in the subject area of information security. *Software engineering*, 2013, № 11, p. 8–16.
6. Stepanov P. A. Automatization of natural language texts processing. *Vestnik Novosib. gos. uni.* Block: Information technology, 2013. Vol. 11, № 2, p. 106–112.
7. Austin J. L. How to do things with words. *Set of lectures*. Oxford: Oxford Univ. Press, 1962.
8. Searle J. R. Speech acts: an essay in the philosophy of language. Cambridge: Cambridge University Press. 1969.
9. Searle J. R. Vandervecken D. Foundations of Illocutionary Logic. Cambridge Univ. Press: Cambridge, 1985.
10. Austin J. L. Word as action. *New in foreign linguistics*. Ed. 17: The theory of speech acts. M., 1986, p. 22–129.
11. Searle J. R. Classification of illocutionary acts. *New in foreign linguistics*. Ed. 17: The theory of speech acts. Moscow, 1986, p. 170–194.
12. Searle J. R. Vanderveken D. Basic concepts of calculus of speech acts. *New in foreign linguistics*. Ed. 18: Logical analysis of natural language. Moscow, 1986, p. 242–264.
13. Vanderveken D. Meaning and Speech Acts. Cambridge Univ. Press, 1990. Vol. 1: Principles of Language Use.
14. Vanderveken D. Meaning and Speech Acts. Cambridge Univ. Press, 1991. Vol. 2: Formal Semantics of Success and Satisfaction.
15. Palchunov D. E. On the logical analysis of natural language theory calculations and specification languages. Novosibirsk, 1995, Vol. 152: Computing systems, p. 61–75.
16. Palchunov D. E. Algebraic description of the meaning of natural language sentences. *Computing systems*, Novosibirsk, 1997. Vol. 158, p. 127–148.
17. Pal'chunov D. E. Algebraische Beschreibung der Bedeutung von Aeusserungen der natuerlichen Sprache // In Zelger, Josef / Martin Maier (Hrsg.): GABEK. Verarbeitung und Darstellung von Wissen. Innsbruck-Wien, STUDIENVerlag, 1999, 310–326.