

В. А. СОБОЛЕВСКИЙ

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук,
Санкт-Петербург

ПРОГРАММНЫЙ КОМПЛЕКС АВТОМАТИЗИРОВАННОЙ ГЕНЕРАЦИИ СЕРВИСОВ НА БАЗЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

В данном докладе представлены исследования, связанные с созданием программного комплекса, реализующего автоматизированную генерацию программ-сервисов для систем искусственного интеллекта на базе искусственных нейронных сетей. Сгенерированные сервисы адаптированы для использования в программных комплексах, реализующих сервис-ориентированный подход и концепцию интернета вещей.

Введение. В современных отраслях науки и техники всё шире применяются системы искусственного интеллекта на базе искусственных нейронных сетей (ИНС). Однако на данный момент для применения данных технологий требуется участие профильных специалистов, которые должны не только создать подобную систему, но и интегрировать её в существующие программные комплексы. С другой стороны, всё большую популярность приобретают сервис-ориентированная архитектура (СОА) и концепция интернета вещей (ИВ), для организации крупных распределённых программных комплексов. При этом программы на базе ИНС часто являются вспомогательными, в рамках подобных крупных программных комплексов. Актуальной становится задача не просто интеграции программ на базе ИНС, а реализация их как автономных сервисов, для использования в рамках СОА и концепции ИВ. Предлагаемый доклад посвящён программе, реализующей автоматизированную генерацию программных оболочек для систем на базе ИНС, для их последующей интеграции в программные комплексы, спроектированные на базе СОА.

Состояние исследований. Несмотря на то, что на данный момент программные комплексы на базе ИНС используются преимущественно для решения специфических задач [1–3], всё чаще их начинают применять для более рутинных задач. Одним из ограничений распространения подобных программ всё ещё остаётся необходимость в привлечение к разработке специалистов в области искусственного интеллекта. Именно для автоматизации решения простых типовых задач и предлагается применять разрабатываемый комплекс. В настоящее время уже существует ряд разработок, призванных автоматизировать и упростить процесс создания программных комплексов на базе ИНС [4–6]. Их наличие показывает актуальность данной задачи, а результаты демонстрируют перспективность подобного подхода. Программный комплекс, представляемый в данном докладе, делает акцент на автоматизацию процесса генерации программных оболочек и интеграцию разработанных ИНС различных архитектур в крупные программные комплексы. Существующие на данный момент решения имеют более узкое применение или используются только для ИНС специфической архитектуры (к примеру, для свёрточных нейронных сетей).

Используемые технологии. СОА подразумевает модульный подход к разработке программного приложения [7]. Данный подход базируется на принципах прозрачности и масштабируемости, что позволяет расширять функциональность программного продукта не модификацией уже существующих модулей, а добавлением новых сервисов. Поэтому, для следования заданной СОА парадигме, требуется, чтобы генерируемые программные сервисы имели типовой программный или веб-интерфейс и использовали распространённые протоколы обмена данными. В разрабатываемом программном комплексе генерируется автономный сервис, содержащий заданную пользователем ИНС, обученную для решения конкретной задачи. Сервис состоит из самой ИНС, а также программных оболочек, реализующих интерфейс REST и SOAP [8–10]. Данный сервис кроссплатформенный и может быть без предварительной установки и настройки дополнительного программного обеспечения запущен на ряде операционных систем (что возможно благодаря кроссплатформенности языка реализации данных модулей – python [11]). А реализация нескольких программных оболочек позволяет обращаться к ИНС как к сервису, через наиболее распространённые веб-интерфейсы. При этом, созданную ИНС всё ещё можно использовать как простую программную библиотеку, обращаясь к ней через соответ-

ствующие программные интерфейсы. Так же, сгенерированные программным комплексом сервисы не обязательно должны быть установлены на одном и том же компьютере. Они могут быть распределены между различными компьютерами или же располагаться в облачных хранилищах. Соответственно, подобные сервисы могут быть использованы в системах поддерживающих как парадигму СОА, так и концепцию ИВ. При этом сам программный комплекс внутри себя так же построен по принципам СОА. Это является необходимым требованием, учитывая, что коммерческое использование программного комплекса потребует больших вычислительных ресурсов. Наиболее простым и удобным в таких условиях будет именно распределённый подход к архитектуре программного комплекса. Дополнительным достоинством СОА является и то, что добавлять новую функциональность к разрабатываемой программе, можно быстро и удобно, без модификации существующего программного кода. В связи с этим, разрабатываемый программный комплекс постоянно растёт, расширяясь как архитектурами ИНС, с которыми он может работать, так и веб-интерфейсами, которые могут поддерживать генерируемые сервисы.

Апробация. Разработанный программный комплекс к настоящему времени уже использовалась для генерации сервисов на базе простых глубоких нейронных сетей, которые были внедрены в сторонние программные продукты для решения конкретных прикладных задач [12–13]. Разнородность решаемых задач демонстрирует относительную универсальность предлагаемого программного комплекса. На рисунке представлена схема программного комплекса «ПРОСТОР» [13], предназначенного для прогнозирования и мониторинга уровня воды в реках. В рамках данного программного комплекса была использована ИНС для прогнозирования уровня воды. Сама ИНС и её обёртка были реализованы с помощью представленного в докладе программного комплекса. Представленная на рисунке схема демонстрирует пример реализации концепции СОА и интеграции программно-сгенерированного сервиса на базе ИНС.

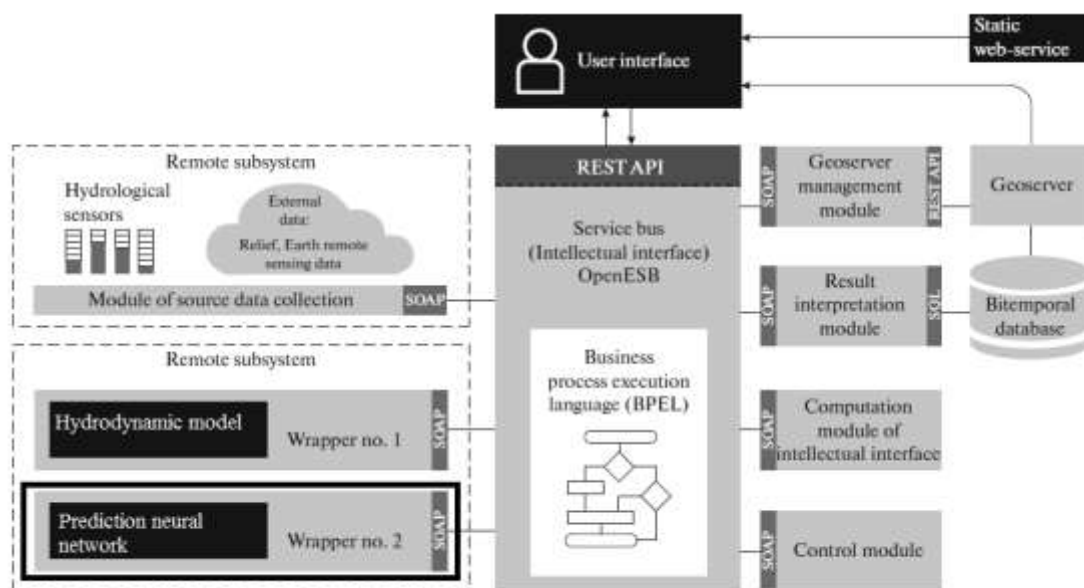


Рисунок. Схема программного комплекса «ПРОСТОР»

Заключение. На данный момент программный комплекс находится на стадии прототипа и используется при разработке ряда сторонних приложений. Для начала полноценной эксплуатации требуется, прежде всего, доработка пользовательского интерфейса приложения. Поскольку на этапе прототипирования продукт используется специалистами в машинном обучении, и поэтому текущий интерфейс не адаптирован для использования обычными пользователями. Из-за этого, на данный момент, не решается одна из основных поставленных перед программным комплексом задач – доступность для широкой аудитории пользователей. Так же, из-за высоких требований к производительности при функционировании данного программного продукта, для коммерческого использования требуется перенос программного комплекса на высокопроизводительные сервера. Специфика вычислений, при обучении ИНС, накладывает ряд требований

на аппаратную часть, а коммерческое использование подразумевает параллельное обучение нескольких моделей, что может существенно нагрузить систему. Несмотря на заложенную в архитектуру программного комплекса параллельность вычислений, требуется провести дополнительные исследования и стресс-тесты, чтобы определить конкретные требования к аппаратной части. Сам программный комплекс так же продолжает расширяться новыми архитектурами искусственных нейронных сетей, для которых могут быть сформированы сервисные оболочки. На данный момент уже реализована и проходит тестирование возможность работы со свёрточными нейронными сетями.

*Работа выполнена при поддержке гранта РФФИ № 19-37-90112
и бюджетной темы 0073- 2019-0004.*

ЛИТЕРАТУРА

1. Ateeq-ur-Rauf, Ghumman A.R., Ahmad S., Hashmi H.N. Performance assessment of artificial neural networks and support vector regression models for stream flow prediction. *Environmental Monitoring and Assessment*. 2018. V. 190, issue 12, article № 704.
2. Alizadeh Z., Yazdi J., Kim J.H., Al-Shamiri A.K. Assessment of machine learning techniques for monthly flow prediction. *Water (Switzerland)*. 2018. V. 10, issue 11, article № 1676.
3. Lantrip J., Griffin M., Aly A. Results of near-term forecasting of surface water supplies. In: *World Water Congress 2005: Impacts of Global Climate Change - Proceedings of the 2005 World Water and Environmental Resources Congress. Anchorage, Alaska, US*. 2005. Pp. 436. Doi: 10.1061/40792(173)447.
4. Geng Z., Wang Y. Automated design of convolutional neural network with multi-scale filters for cost-efficient seismic data classification. *Nature Communications*. 2020. V. 11, Issue 1.
5. Witsuba M., Rawat A., Pedapati T. Automation of deep learning. *Proceedings of the 2020 International Conference on Multimedia Retrieval, ICMR 2020*. 2020. P. 5-6.
6. Baker B., Gupta O., Nail N., Raskar R. Designing neural network architectures using reinforcement learning. *5th International Conference on Learning Representations, ICLR 2017*. 2017.
7. **Bell M.** Introduction to Service-Oriented Modeling, in *Service-Oriented Modeling^ Service Analysis, Design and Architecture*. Wiley & Sons, 2008, pp. 390. ISBN 978-0-470-14111-3.
8. Mesmoudi Y., Lamnaour M., Khamlichi Y.E.L., Tahiri A., Touhafi A., Braeken A. Design and implementation of a smart gateway for IoT applications using heterogeneous smart objects. *4th International Conference on Cloud Computing Technologies and Applications, Cloudtech*. 2018.
9. **Hanes D.** IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things. USA: Cisco Press, 2017, pp. 576.
10. **Erl T.** Service-Oriented Architecture: Analysis and Design for Services and Microservices (2nd Edition). USA: Prentice Hall, 2016, pp. 416.
11. **John V. Guttag** Introduction to Computation and Programming Using Python: With Application to Understanding Data (2nd Edition). USA: MIT Press, 2016, pp. 472. ISBN 978-0-262-52962-4.
12. Mikhailov V., Specivtsev A., Sobolevsky V., Kartashev N. Multi-Model Estimation of the Dynamics of Plant Community Phytomass. *The 13th IEEE International Conference Application of Information and Communication Technologies (AICT)*. 2019. P. 324-328.
13. Zelentsov V.A., Alabyan A.M., Krylenko I.N., Pimanov I.Yu., Ponomarenko M.R., Potryasaev S.A., Semenov A.E., Sobolevskii V.A., Sokolov B.V., Yusupov R.M. A Model-Oriented System for Operational Forecasting of River Floods. *Herald of the Russian Academy of Sciences*. 2019. Volume 89, Issue 4. P. 405-417. DOI: 10.1134/S1019331619040130.

V.A. Sobolevskii (St. Petersburg Institute for Informatics and Automation, Russian Academy of Sciences, St. Petersburg)

Software Complex for Automated Generation of Services Based on Artificial Neural Networks

This report presents research related to the creation of a software package that implements the automated generation of service programs for artificial intelligence systems based on artificial neural networks. The generated services are adapted for use in software systems that implement a service-oriented approach and the concept of the Internet of Things.