

**АВТОМАТИЗИРОВАННЫЙ МЕТОД ПОДГОТОВКИ ДАННЫХ ДЛЯ МОДЕЛИ МАШИННОГО
ОБУЧЕНИЯ ПО ДЕТЕКТИРОВАНИЮ ГРАНИЦЫ АВРОРАЛЬНОГО ОВАЛА ПО РАЗРЯЖЕННОЙ СЕТИ
ГНСС СТАНЦИЙ**

**Борвенко Е.А. (Иркутский государственный университет)
Научный руководитель – Зав. Лабораторией Робототехники и программирования
ИГУ, научный сотрудник ИСЗФ СО РАН - Веснин А.М. (ИГУ)**

Введение. Авроральный овал является динамичной структурой как с точки зрения уровня возмущенности внутри овала, так и с точки зрения определения его границы [1, 2]. Высокий уровень возмущенности может негативно влиять на ГНСС измерения и приводить к понижению точности определения координат. При этом важно определять границу овала с точки зрения ионосферных возмущенности, которые влияют на сигналы ГНСС, используя их в качестве зондирующих сигналов. Определение границ аврорального овала как непрерывной структуры по данным ГНСС-станций представляет собой сложную задачу из-за разреженности сети наземных приемников широком диапазоне долгот. Высокая изменчивость ионосферы, не позволяет делать пространственную интерполяцию. В данной работе предложен автоматизированный метод подготовки данных, направленный на создание датасета для обучения нейронной сети, которая сможет работать в условиях разреженной сети.

Основная часть. Основная цель — выявить моменты времени пересечения спутниками границы аврорального овала, используя индекс ионосферной возмущенности ROTI (Rate of Total Electron Content Index). Было проверено две гипотезы:

- точка пересечения может быть использованы сами временные ряды, где критерием пересечения может изменения статистических параметров ряда данных;
- точка пересечения определяется на ограниченных территориях, где существует хорошее покрытие данных.

Для проверки гипотезы о возможности автоматического детектирования пересечения границы овала по временным рядам была проведена серия расчетов. Исходные данные ГНСС были получены из системы SIMuRG[3], после чего были отобраны станции, расположенные вблизи предполагаемой границы аврорального овала. Для каждого спутника строились два графика: первый отражал изменение ROTI во времени, а второй специально разработанный индекс для детектирования пересечения границы.

Для расчета индекса применялся метод скользящих окон, где два окна проходили все точки ряда, последнее значение первого окна совпадало с первым значением второго. Для каждого окна вычислялось среднеквадратическое отклонение (СКО) и медиана (два варианта расчета индекса), после чего рассчитывалось их отношение. Предполагалось, что внутри овала значения ROTI выше, чем снаружи, и при пересечении границы наблюдается резкий рост этого индекса. В результате можно было определить точки входа и выхода спутника из овала по экстремумам на графике с индексом. Тем не менее, возникли две проблемы. Первая заключалась в выборе оптимального размера скользящего окна, поскольку он оказывал значительное влияние на точность определения точки пересечения. Вторая проблема связана с тем, что не все спутники фактически пересекают границу овала, что требовало дополнительной фильтрации данных. Для этого данные спутников разделялись на две части, и анализировалось отношение их медианных значений. Если оно превышало заданный порог, спутник считался пересекавшим границу, иначе — нет. Однако такой метод фильтрации приводил к тому, что около 20% отобранных данных оказывались некорректными. Увеличение порогового значения снижало процент ошибочных данных, но одновременно уменьшало объем датасета, что требовало значительного увеличения объема обработанных данных.

Для решения этих проблем была предложена альтернативная гипотеза для метода разметки границы овала на временных рядах. На основе значений ROTI полученных в регионах с

хорошим пространственным покрытием данных создавалась карта, где каждой точке с определенной широтой и долготой, соответствующих подионосферной точке, присваивалось значение индекса. Затем применялось скользящее окно фиксированного размера, в пределах которого вычислялось медианное значение ROTI, что позволило прийти от нерегулярного распределения данных к данным распределенным на регулярной сетке. Это позволило снизить влияние выбросов и улучшить качество определения границы. Для исключения возможных белых пятен отсутствующие значения были получены с помощью интерполяции существующих данных, а граница определялась по заданному уровню индекса (0.07 TECu/minute). Для установки точки пересечения границы овала на временных рядах использовался следующий подход: по серии карт и найденных на них границах определяем сколько раз траектория спутника пересекла эту границу. Траектория может находиться полностью снаружи или полностью внутри области овала, тогда не будет ни одного пересечения. Так же траектория может пересекать границу один раз. И последний случай, в силу динамической природы границы, пересечений может быть множество. На основе этих случаев формируется датасет, который может быть использован для обучения модели, на вход которой могут быть поданы отдельные временные ряды, а на выходе будут получены вероятности присутствия в данных случаев описанных выше.

Выводы. Результатом работы стал метод, который ранее не применялся для разметки данных для получения датасета пригодного для моделей машинного обучения. Данный подход позволяет автоматизировать процесс построения границы аврорального овала по данным ГНСС без необходимости использования дополнительных картографических данных. Был получен датасет с разметкой границы аврорального овала, который может быть использован для обучения нейросетевой модели.

Список использованных источников:

1. Edemskiy IK, Yasyukevich YV. Auroral oval boundary dynamics on the nature of geomagnetic storm. Remote Sensing. 2022 Oct 31;14(21):5486.
2. Vasiliev A, Garashchenko A, Sidorov D, Yasyukevich Y, Edemskiy I, Vesnin A. Computer vision for GNSS-based detection of the auroral oval boundary. International Journal of Artificial Intelligence. 2021;19(2):132-51.
3. Yasyukevich YV, Kiselev AV, Zhivetiev IV, Edemskiy IK, Syrovatskii SV, Maletckii BM, Vesnin AM. SIMuRG: System for ionosphere monitoring and research from GNSS. GPS solutions. 2020 Jul;24:pp.1-12.