

Применение алгоритмов сжатия информации для задачи когнитивного отображения проектного пространства многослойных экранов защиты от радиации

Л.А. Зинченко ¹✉, В.В. Казаков ¹, Б.В. Карышев ¹

¹ МГТУ им. Н.Э. Баумана, г. Москва, 105005, Россия

Ссылка для цитирования

Зинченко Л.А., Казаков В.В., Карышев Б.В. Применение алгоритмов сжатия информации для задачи когнитивного отображения проектного пространства многослойных экранов защиты от радиации // Программные продукты и системы. 2024. Т. 37. № 2. С. 178–185. doi: 10.15827/0236-235X.142.178-185

Информация о статье

Группа специальностей ВАК: 2.3.7

Поступила в редакцию: 31.01.2023

После доработки: 14.11.2023

Принята к публикации: 21.11.2023

Аннотация. В статье рассмотрены особенности применения алгоритмов снижения размерности в задачах проектирования многослойных радиационно стойких экранов защиты электронной аппаратуры для использования в космическом пространстве. Алгоритмы проецируют исследуемые данные о многослойных защитных экранах из пространства с высокой размерностью в пространство с низкой размерностью. При этом сохраняется семантика данных, что позволяет визуализировать большие наборы информации высокой размерности и многократно упрощает пользовательский визуальный анализ, а также применение некоторых алгоритмов и подходов в автоматизированном режиме. В статье рассмотрено применение следующих алгоритмов снижения размерности: анализ главных компонент (PCA), анализ ядра главных компонент (KernelPCA), стохастическое вложение соседей с t-распределением (t-SNE), равномерное приближение и проекция (UMAP), автоэнкодер (AE), вариационный автоэнкодер (VAE). В случае нейросетевых архитектур сжатия представлены предлагаемые архитектуры сетей, используемые в расчетах и тестировании. По предложенной методике исследована целесообразность комбинации нескольких алгоритмов снижения размерности, применяемых по цепочке. На основании проведенного исследования сделан вывод о результативности применения указанных алгоритмов, а также их комбинации для последующей обработки или визуализации. Кратко рассмотрено программное обеспечение, реализующее один из предложенных подходов в анализе и обработке информации о многослойных радиационно стойких экранах защиты электронной аппаратуры для использования в космическом пространстве. В результате сделан выбор в пользу алгоритма UMAP. Для анализа конфигураций с достаточно большим количеством параметров рекомендуется использовать алгоритм t-SNE с предварительным сжатием алгоритмом UMAP, который упрощает исходный набор данных, улучшая результат работы t-SNE.

Ключевые слова: когнитивная визуализация, сжатие данных, алгоритмы снижения размерности, проецирование

Введение. Для обеспечения стабильного функционирования электронной аппаратуры в космическом пространстве применяются экраны защиты от радиации. При их проектировании необходимо сравнить большое количество альтернативных проектных решений и выбрать наиболее подходящее [1]. Вычисление характеристик каждого альтернативного проектного решения требует больших временных затрат. При этом проектировщику необходимо провести анализ полученных решений. В существующих программных средствах, таких как SRIM [2] и GEANT4 [3], анализ возможных структур экранов защиты не реализован. Таким образом, возникает необходимость в ПО, упрощающем выбор проектного решения.

Для решения проблемы анализа конфигураций экранов защиты предложена методика с использованием алгоритма t-SNE [4]. Производится декомпозиция параметров, характеризующих структурные составляющие многослой-

ного защитного экрана, на различные уровни – многослойный экран, материалы и отдельные химические элементы. После чего полученные параметры представляются в численном виде. В случае с характеристиками, не являющимися численными, можно использовать кодировку, где каждому возможному значению параметра соответствует свой уникальный численный код, либо так называемый метод one-hot encoding, при котором каждому возможному значению характеристики соответствует элемент множества, принимающий значение 1, если значение характеристики соответствует данному элементу, и 0, если не соответствует. Таким образом, каждой конфигурации соответствует множество численных значений, которые можно рассматривать как точки в пространстве высокой размерности. Далее используется алгоритм t-SNE, проецирующий многомерные данные в двухмерное пространство с сохранением общей структуры данных. После этого строятся

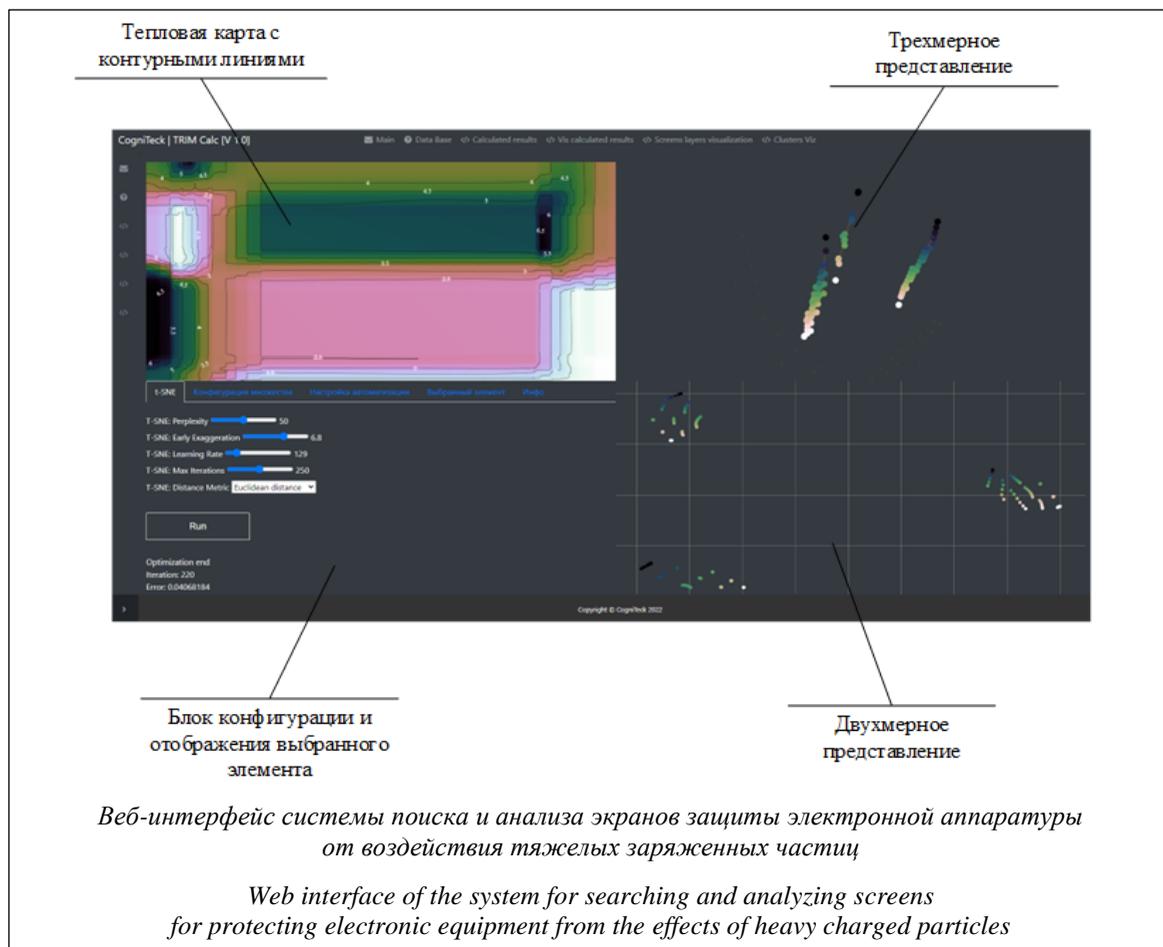
когнитивные визуальные образы на основе полученной проекции точек, отражающих возможные конфигурации, благодаря чему значительно упрощается восприятие данных и разработчик может определить степень сходства конфигураций между собой [5]. Такой подход позволяет приблизительно оценить характеристики группы конфигураций многослойных экранов защиты, проведя расчеты для одной или нескольких из них.

Важным элементом системы является методика сжатия данных, которая напрямую влияет на качество получаемых когнитивных представлений. В случае последующего прогноза некоторых целевых показателей для оптимизации потребления памяти, нагрузки на центральный процессор или графического ядро (в зависимости от алгоритма прогноза и его реализации) сжатие может выступать дополнительным этапом предобработки, повышающим эффективность последующего этапа прогноза. В таком случае выбор наилучшего алгоритма или предлагаемой новой методики, включающей несколько стадий сжатия различными алгоритмами, является крайне важным этапом

при проектировании систем визуализации и прогноза, в связи с чем необходимо провести анализ и оценку алгоритмов сжатия, а также методологии многостадийного сжатия информации для выбора наилучшего решения.

Интерфейс системы анализа конфигураций с когнитивными элементами, изображенный на рисунке, разделен на четыре зоны, три из которых представляют когнитивно-графический образ проекции многомерного множества параметров конфигураций в пространство с низкой размерностью, а один – настройки параметров алгоритма сжатия и графического образа, а также задание исследуемого набора конфигураций экранов защиты. Тепловая карта показывает двухмерное цветовое представление одного из исходных параметров и трехмерное представление, отражающее значение определенного параметра конфигурации, вынося его на ось z. В двухмерном и трехмерном представлениях также цветом выделен один из параметров конфигурации (в данном случае – масса).

Основная задача заключается в том, чтобы оценить различные алгоритмы сжатия много-



мерных данных и их комбинаций и найти оптимальный вариант, позволяющий добиться наименьших потерь информации для последующих задач прогнозирования и визуализации.

В разработанной системе на вход подаются параметры каждого защитного экрана. Параметры химического элемента в составе слоя: процент содержания химических элементов в материале (для составных материалов, например, сплав и другие), порядковый номер химического элемента в периодической таблице, период химического элемента в периодической таблице, группа химического элемента в периодической таблице, подгруппа химического элемента в периодической таблице, атомная масса химического элемента, семейство химического элемента, электроотрицательность химического элемента, температура плавления химического элемента, температура кипения химического элемента, вес материала на 1 м³, плотность химического элемента.

Параметры слоя многослойного экрана: вес 1 м³ материала (слоя многослойного экрана), общий вес слоя, количество химических элементов, из которых состоит материал (слоя многослойного экрана), толщина слоя многослойного экрана.

Параметры многослойного экрана в целом: количество слоев в данном многослойном экране, вес многослойного экрана площадью 1 м² (толщина экрана определяется исходя из параметров его слоев), общая толщина многослойного экрана (в мм).

Часть из приведенных параметров используется напрямую в структуре данных для непосредственного сжатия, а у некоторых (таких, как порядковый номер в периодической таблице) необходимо корректно сопоставить данные о химическом элементе и его объекте в рамках разработанной системы.

Предложенный подход работы с многомерными данными с применением алгоритма t-SNE был реализован в программном комплексе CogShield [6]. Для дальнейшего повышения качества визуального анализа многомерных данных необходимо проанализировать иные алгоритмы сжатия и их комбинации и предложить подход в оценке алгоритмов, не зависящий от задачи приложения.

Обзор алгоритмов снижения размерности

Существует большое количество алгоритмов снижения размерности (см. [6]), которые можно комбинировать между собой. Выбор ал-

горитма или их комбинации напрямую влияет на результативность описанной выше методики проектирования экрана защиты от радиации. При этом важно выбрать для использования такой алгоритм, который максимально сохраняет полезную информацию в данных и исключает дублирование информации или малозначимые компоненты.

Автоэнкодер (Autoencoder – AE) представляет две связанные между собой нейронные сети: кодер, сжимающий данные до заданной размерности, и декодер, разжимающий их обратно [7]. Во время обучения их объединяют в одну нейросеть, задача которой сжать вектор, а потом восстановить его как можно точнее. Чаще всего обучение происходит методом градиентного спуска. Существенным недостатком автоэнкодера является его низкая вычислительная эффективность.

Вариационный автоэнкодер (Variational Autoencoder – VAE) позволяет сформировать сжатое множество в соответствии с заданным законом распределения (обычно используется нормальное распределение) [8]. Это приводит к тому, что вариационный декодер преобразует точки в окрестности, полученные энкодером, в изображения, похожие на исходные. Благодаря этому свойству вариационный автоэнкодер используется для генерации данных, схожих с элементами некоего множества. При исследовании выбрана одинаковая структура вариационного автоэнкодера и автоэнкодера. В качестве функции активации для нейронных сетей использована ReLU [9], так как она позволяет автоэнкодерам выявлять нелинейные закономерности в данных. В эксперименте энкодер и декодер включали по пять слоев, причем количество нейронов зависело от размера входного вектора. В таблице 1 представлено количество нейронов в каждом слое кодера. Декодер имел симметричную структуру.

Таблица 1

Структура кодера

Table 1

Coder structure

Слой				
1	2	3	4	5
Входной вектор размера n	$2n$	n	n	Вектор скрытого состояния (в данной работе размер равен 2)

Метод анализа главных компонент (PCA) состоит в выборе главных компонент – новой системы координат, лучше описываю-

щей данные [10]. В случае с нормальным распределением данные напоминают многомерный эллипс, главные компоненты совпадают с осями этого эллипса. Затем вычисляются дисперсии значений по компонентам. Далее компонента, вносящая минимальный вклад в информативность, удаляется.

Метод KernelPCA представляет собой улучшенную версию PCA, способную выявлять нелинейные зависимости в данных [11].

Алгоритм t-SNE [12] включает следующие шаги. Для всех элементов множества вводится и вычисляется функция, преобразующая многомерное евклидово расстояние между точками в условные вероятности, отражающие сходство точек между собой. Элементы исходного множества случайным образом проецируются в пространство низкой размерности, и для каждого элемента нового множества вычисляется аналогичный показатель. Посредством изменения положения элементов множества-проекции увеличивается сходство структуры исходного множества и его проекции. Сходство оценивается по сумме дивергенций Кульбака-Лейблера [13, 14] для всех точек. У алгоритма t-SNE есть очень важный параметр – перплексия. При ее высоком значении сохраняется глобальная структура данных, а при низком – локальная.

Алгоритм UMAP [15] по принципу работы похож на алгоритм t-SNE. Его преимуществами являются улучшенное быстродействие, сохранение глобальной структуры множества при проецировании: расстояние между кластерами после сжатия t-SNE ничего не говорит о схожести кластеров между собой, в то время как после алгоритма UMAP близость кластеров свидетельствует об их схожести. Алгоритм UMAP в отличие от алгоритма t-SNE можно эффективно применять не только для отображения данных, но и для общего снижения размерности с целью избавления от незначимых компонент и мультиколлинеарности в данных.

Сравнение алгоритмов и их комбинаций

Для количественного определения качества работы алгоритма снижения размерности в работе использованы алгоритмы классификации, для оценки результативности которых была использована метрика Ассигасу, вычисляемая как отношение количества правильно распознанных элементов к их общему количеству. Сначала алгоритм классификации обучался на исходном наборе данных, а после – на сжатом.

Разность между результатом после сжатия и до сжатия принята как показатель качества работы алгоритма сжатия. Чем он выше, тем меньше информации было потеряно при сжатии. Применялись следующие алгоритмы: градиентный бустинг, k-средних и k-ближайших соседей. Если точность прогноза алгоритма после сжатия падала ниже вероятности случайного угадывания, то в качестве результата принималось число –1. При иных значениях для неудачного сжатия результаты менялись.

Для исследования были выбраны следующие наборы данных:

- случайная выборка из 1 000 элементов из набора MNIST (<http://yann.lecun.com/exdb/mnist/>), представляющего собой набор черных белых рукописных цифр 28×28 пикселей, подаваемых на вход в виде числовых массивов; целевое значение для прогноза – значение цифры;
- Water Quality (<https://www.kaggle.com/datasets/adityakadiwal/water-potability>) – набор химических и физических характеристик воды; целевое значение для прогноза – пригодность воды для питья;
- Mobile Price Classification (<https://www.kaggle.com/datasets/iabhishekoofficial/mobile-price-classification>) – набор характеристик телефонов; целевое значение для прогноза – стоимостная группа;
- Titanic (<https://www.kaggle.com/c/titanic>) – информация о пассажирах; целевое значение для прогноза – выжил ли человек после крушения;
- Iris (<https://www.kaggle.com/datasets/uciml/iris>) – набор параметров цветков ириса; целевое значение для прогноза – подвид;
- Student's scores (<https://www.kaggle.com/datasets/samarsaeedkhan/scores>) – набор оценок студентов по различным предметам; целевое значение для прогноза – удовлетворенность оценками.

В ходе эксперимента алгоритмы UMAP и t-SNE использовались с различными значениями количества соседей и переплексии соответственно, после чего выбирался наилучший результат. Параметры остальных алгоритмов в ходе эксперимента не менялись.

В процессе работы также была исследована результативность предварительного сжатия с использованием алгоритмов PCA, KernelPCA и UMAP. Остальные алгоритмы из перечисленных выше для многоэтапного сжатия не подходят, так как алгоритм t-SNE не способен сжи-

мать данные в размерности больше 3, а автоэнкодеры обладают слишком низкой вычислительной эффективностью. Три последних набора данных обладают относительно низкой исходной размерностью (3–6 измерений), поэтому исследовать на них предварительное сжатие не имело смысла, поскольку все отображение будет выполнено первым примененным алгоритмом.

Для эксперимента использовались следующие реализации алгоритмов:

- t-SNE – пакет sklearn (<https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html>);
- UMAP – реализация авторов алгоритма (<https://umap-learn.readthedocs.io/en/latest/index.html>);
- PCA – пакет sklearn (<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>);
- KernelPCA – пакет sklearn (<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.KernelPCA.html>);
- AE и VAE – реализованы с использованием keras (<https://keras.io/>).

Для выполнения эксперимента был использован шестиядерный процессор AMD Ryzen 5 5500U с частотой 2,1 ГГц и 16 Гб оперативной памяти. Результаты эксперимента представлены в таблицах 2 и 3.

Таблица 2

Результаты работы алгоритмов без предварительного сжатия

Table 2

Results of algorithms without precompression

Метод	Количество корректно сжатых наборов данных (из 6)	Изменение среднего значения метрики Accuracy
t-SNE	6	0,0497
UMAP	6	0,0454
KernelPCA	5	-0,0942
PCA	3	-0,1789
AE	3	-0,2030
VAE	1	-0,3886

Анализ результатов предлагаемого подхода оценки алгоритмов сжатия в таблицах 2 и 3 позволяет сделать вывод, что при проецировании лучше всего сохраняют целевое значение алгоритмы UMAP и t-SNE. Предварительное сжатие позволяет немного улучшить результат. В работе [16] описано многостадийное сжатие,

но нет описания метода поэтапного сжатия UMAP_и_t-SNE и других приведенных в данной работе. Наилучший результат дает предварительное сжатие алгоритмом UMAP с последующим сжатием t-SNE.

Таблица 3

Результаты работы алгоритмов с предварительным сжатием

Table 3

Results of algorithms with precompression

Метод	Количество корректно сжатых наборов данных (из 3)	Изменение среднего Accuracy
UMAP_и_t-SNE	3	0,0044
UMAP_и_PCA	3	-0,0054
KernelPCA_и_UMAP	3	-0,0080
PCA_и_t-SNE	3	-0,0089
UMAP	3	-0,0122
t-SNE	3	-0,0149
KernelPCA_и_t-SNE	3	-0,0337
KernelPCA_и_PCA	3	-0,0886
KernelPCA_и_AE	3	-0,1090
UMAP_и_KernelPCA	2	-0,1156
PCA_и_UMAP	2	-0,1212
UMAP_и_AE	2	-0,1633
AE	2	-0,1904
KernelPCA	2	-0,2098
PCA_и_AE	2	-0,2099
KernelPCA_и_VAE	2	-0,2309
PCA_и_KernelPCA	2	-0,2392
UMAP_и_VAE	2	-0,2733
PCA_и_VAE	2	-0,2738
PCA	1	-0,3149
VAE	0	-0,4004

Результаты показывают, что в сравнении с современными аналогами (например, KernelPCA) и подходами в сжатии данных алгоритмы UMAP и t-SNE, а также предлагаемая методика двухэтапного сжатия (метод UMAP_и_t-SNE из таблицы 3) позволяют получить в среднем более высокие результаты прогноза целевых значений в исследуемых тестовых датасетах, в том числе и по сравнению с методом прогноза без сжатия, что является важным наблюдением, которое может позволить не только оптимизировать использование памяти и загрузки процессора при самом прогнозе целевых значений, но и поднять качество прогноза необходимых показателей исследуемых

объектов. Интуиция дальнейшего применения полученных результатов в задачах когнитивной визуализации многомерных данных заключается в том, что алгоритм сжатия, позволяющий получить более точный прогноз целевой метрики методами машинного обучения при построении когнитивного образа для визуальной аналитики экспертом, позволит выявить больше закономерностей и инсайтов в процессе взаимодействия эксперта и системы анализа.

Рассматривая полученные результаты применительно к разработанной системе анализа многослойных радиационно стойких защитных экранов, можно сделать вывод о том, что необходимо добавить возможность сжатия методом UMAP и t-SNE, показавшим наилучшие результаты, и сделать его сжатием по умолчанию, а также добавить сжатие методом UMAP с возможностью переключения между алгоритмами, чтобы дать возможность эксперту системы посмотреть на данные в разных сжатиях и выбрать наиболее подходящее с точки зрения эксперта представление.

Заключение

В статье предложена методика сравнения алгоритмов сжатия для последующей визуализации многомерных данных. По предложенной методике было выполнено сравнение алгоритмов t-SNE, UMAP, PCA, KernelPCA, AE, VAE. Также исследована эффективность предварительного сжатия данных и двухэтапного применения алгоритмов снижения размерности.

При построении когнитивных образов для последующего визуального анализа пространства решений рекомендуется использовать алгоритм UMAP, поскольку он сохраняет целевое значение так же хорошо, как и t-SNE, то есть позволяет проводить анализ с большей точностью, чем другие алгоритмы, но требует меньших вычислительных затрат. Для анализа конфигураций с достаточно большим количеством параметров рекомендуется использовать алгоритм t-SNE с предварительным сжатием алгоритмом UMAP, который упрощает исходный набор данных, тем самым улучшая результат работы t-SNE.

Список литературы

1. Терехов В.В., Михайличенко С.С. Методика определения толщины экрана для защиты микроэлектромеханических систем от воздействия высокоэнергичных ионов // Наукоемкие технологии и интеллектуальные системы: матер. Междунар. науч.-технич. конф. 2016. С. 356–358.
2. Mitsi E., Koutsomitis K., Apostolopoulos G. On the use of SRIM for calculating arc-dpa exposure. ArXiv, 2023, art. 2307.12867. URL: <https://arxiv.org/abs/2307.12867> (дата обращения: 12.11.2023).
3. Sytov A., Bandiera L., Cho K. et al. Geant4 simulation model of electromagnetic processes in oriented crystals for accelerator physics. J. Korean Phys. Soc., 2023, vol. 83, pp. 132–139 (дата обращения: 12.11.2023). doi: 10.1007/S40042-023-00834-6.
4. Шахнов В.А., Казаков В.В., Глушко А.А. и др. О возможности применения алгоритмов эволюционных вычислений при проектировании экранов радиационной защиты // Радиолокация и связь – перспективные технологии: XVII науч.-технич. конф. 2019. С. 87–90.
5. Казаков В.В. Генерация когнитивного представления информации о многослойных радиационно-стойких защитных экранах // ИС&ИТ: сб. тр. конф. 2022. Т. 1. С. 187–194.
6. Waggoner Ph.D. Modern Dimension Reduction (Elements in Quantitative and Computational Methods for the Social Sciences). Cambridge University Press, 2021, 98 p.
7. Meneghetti L., Demo N., Rozza G. A dimensionality reduction approach for convolutional neural networks. Appl. Intell., 2023, vol. 53, pp. 22818–22833. doi: 10.1007/s10489-023-04730-1.
8. Гергет О.М., Лаптев В.В. Разработка моделей искусственных нейронных сетей для вариационного автоэнкодера // МСИТ: матер. XVIII Междунар. науч.-технич. конф. 2021. С. 51–52.
9. Соснин А.С., Сулова И.А. Функции активации нейросети: сигмоида, линейная, ступенчатая, RELU, TANH // Наука. Информатизация. Технологии. Образование: матер. XII Междунар. науч.-практ. конф. 2019. С. 237–246.
10. Bilokon P., Finkelstein D. Iterated and exponentially weighted moving principal component analysis. SSRN, 2021. URL: <https://ssrn.com/abstract=3913940> (дата обращения: 12.11.2023). doi: 10.2139/ssrn.3913940.
11. Avella-Medina M., Davis R.A., Samorodnitsky G. Kernel PCA for multivariate extremes. ArXiv, 2022, art. 1207.3538. URL: <https://arxiv.org/pdf/2211.13172.pdf> (дата обращения: 12.11.2023).
12. Cai T.T., Ma R. Theoretical foundations of t-SNE for visualizing high-dimensional clustered data. JMLR, 2022, vol. 23, pp. 1–54.
13. Bonnici V. Kullback-Leibler divergence between quantum distributions, and its upper-bound. ArXiv, 2020, art. 2008.05932. URL: <https://arxiv.org/abs/2008.05932> (дата обращения: 12.11.2023).
14. Pain J.-C. Kullback-Leibler divergence for the Fréchet extreme-value distribution. ArXiv, 2023, art. 2303.13153. URL: <https://arxiv.org/pdf/2303.13153.pdf> (дата обращения: 12.11.2023).

15. McInnes L., Healy J., Saul N., Großberger L. UMAP: Uniform manifold approximation and projection. *J. of Open Source Software*, 2018, vol. 3, no. 29, art. 861. doi: 10.21105/joss.00861.

16. Dimple, Misra K., Arun K.G. Evidence for two distinct populations of Kilonova-associated gamma-ray bursts. *ApJL*, 2023, vol. 949, no. 2, art. L22. doi: 10.3847/2041-8213/acd4c4.

Software & Systems

doi: 10.15827/0236-235X.142.178-185

2024, 37(2), pp. 178–185

The problem of cognitive representation of the project space of multilayer radiation shields: Information compression algorithms

Ludmila A. Zinchenko ¹✉, Vadim V. Kazakov ¹, Boris V. Karyshev ¹

¹ Bauman Moscow State Technical University, Moscow, 105005, Russian Federation

For citation

Zinchenko, L.A., Kazakov, V.V., Karyshev, B.V. (2024) 'The problem of cognitive representation of the project space of multilayer radiation shields: Information compression algorithms', *Software & Systems*, 37(2), pp. 178–185 (in Russ.). doi: 10.15827/0236-235X.142.178-185

Article info

Received: 31.01.2023

After revision: 14.11.2023

Accepted: 21.11.2023

Abstract. The paper focuses on dimensional reduction algorithm application in the problems of designing multilayer radiation shields for protecting electronic equipment in outer space. The considered algorithms project the studied data on multilayer shields from high-dimensional space into low-dimensional space with preservation of data semantics, which allows visualizing large sets of high-dimensional information and simplifies user visual analysis, as well as application of some algorithms and approaches in an automated mode. The paper analyses the application of the following dimensionality reduction algorithms: principal component analysis (PCA), kernel principal component analysis (KernelPCA), stochastic neighbour embedding with t-distribution (t-SNE), uniform approximation and projection (UMAP), autoencoder (AE), variational autoencoder (VAE). In terms of neural network compression architectures, the paper presents network architectures used in computation and testing. Moreover, according to the proposed methodology, the authors of the paper investigate feasibility of combining several dimensionality reduction algorithms applied along the chain. Based on the conducted research the authors make a conclusion about the efficiency of the mentioned algorithms, as well as their combination for further processing or visualization. There is a brief description of the software that implements one of the proposed approaches in analyzing and processing information about multilayer radiation shields for electronic equipment used in outer space. On the basis of the conducted research it is recommended to use the UMAP algorithm. To analyze configurations with a sufficiently large number of parameters, it is recommended to use the t-SNE algorithm with precompression by the UMAP algorithm, which simplifies the initial data set, thus improving the result of t-SNE.

Keywords: cognitive visualization, data compression, dimensionality reduction, projection

References

1. Terekhov, V.V., Mikhaylichenko, S.S (2016) 'Method for modeling the thickness of shield for protection microelectromechanical systems from the impact of high-energy ions', *Proc. Int. Sci-Tech. Conf. Tech. & Systems*, pp. 356–358 (in Russ.).
2. Mitsi, E., Koutsomitis, K., Apostolopoulos, G. (2023) 'On the use of SRIM for calculating arc-dpa exposure', *ArXiv*, art. 2307.12867, available at: <https://arxiv.org/abs/2307.12867> (accessed November 12, 2023).
3. Sytov, A., Bandiera, L., Cho, K. et al. (2023) 'Geant4 simulation model of electromagnetic processes in oriented crystals for accelerator physics', *J. Korean Phys. Soc.*, 83, pp. 132–139. doi: 10.1007/S40042-023-00834-6. (accessed November 12, 2023).
4. Shakhnov, V.A., Kazakov, V.V., Glushko, A.A. et al. (2019) 'About possible application of evolutionary computing algorithms for designing radiation protection screens', *Proc. XVII Sci-Tech. Conf. Radar and Communications – Promising Tech.*, pp. 87–90 (in Russ.).
5. Kazakov, V.V. (2022) 'Generation of cognitive representation of information about multilayer radiation-resistant protective screens', *Proc. IS&IT*, 1, pp. 187–194 (in Russ.).
6. Waggoner, Ph.D. (2021) *Modern Dimension Reduction (Elements in Quantitative and Computational Methods for the Social Sciences)*. Cambridge University Press, 98 p.

7. Meneghetti, L., Demo, N., Rozza, G. (2023) 'A dimensionality reduction approach for convolutional neural networks', *Appl. Intell.*, 53, pp. 22818–22833. doi: 10.1007/s10489-023-04730-1.
8. Gerget, O.M., Laptev, V.V. (2021) 'Development of models of artificial neural networks of variational autoencoder', *Proc. Int. Sci-Tech. Conf. YMIT*, pp. 51–52 (in Russ.).
9. Sosnin, A.S., Suslova, I.A. (2019) 'Functions of neural net activation: SIGMOID, LINEAR, STEP, RELU, TANH', *Proc. Sci-Tech. Conf. The Sci. Informatization. Tech. Education*, pp. 237–246 (in Russ.).
10. Bilokon, P., Finkelstein, D. (2021) 'Iterated and exponentially weighted moving principal component analysis', *SSRN*, available at: <https://ssrn.com/abstract=3913940> (accessed November 12, 2023). doi: 10.2139/ssrn.3913940.
11. Avella-Medina, M., Davis, R.A., Samorodnitsky, G. (2022) 'Kernel PCA for multivariate extremes', *ArXiv*, art. 1207.3538, available at: <https://arxiv.org/pdf/2211.13172.pdf> (accessed November 12, 2023).
12. Cai, T.T., Ma, R. (2022) 'Theoretical foundations of T-SNE for visualizing high-dimensional clustered data', *JMLR*, 23, pp. 1–54.
13. Bonnici, V. (2020) 'Kullback-Leibler divergence between quantum distributions, and its upper-bound', *ArXiv*, art. 2008.05932, available at: <https://arxiv.org/abs/2008.05932> (accessed November 12, 2023).
14. Pain, J.-C. (2023) 'Kullback-Leibler divergence for the Fréchet extreme-value distribution', *ArXiv*, art. 2303.13153, available at: <https://arxiv.org/pdf/2303.13153.pdf> (accessed November 12, 2023).
15. McInnes, L., Healy, J., Saul, N., Großberger, L. (2018) 'UMAP: Uniform manifold approximation and projection', *J. of Open Source Software*, 3(29), art. 861. doi: 10.21105/joss.00861.
16. Dimple, Misra, K., Arun, K.G. (2023) 'Evidence for two distinct populations of Kilonova-associated gamma-ray bursts', *ApJL*, 949(2), art. L22. doi: 10.3847/2041-8213/acd4c4.

Авторы

Зинченко Людмила Анатольевна¹, д.т.н.,
профессор, lyudmillaa@mail.ru
Казakov Вадим Вячеславович¹,
аспирант, kazakov-vadim-k@yandex.ru
Карышев Борис Владимирович¹,
студент, boris.karyshev@gmail.com

¹ МГТУ им. Н.Э. Баумана,
г. Москва, 105005, Россия

Authors

Ludmila A. Zinchenko¹, Dr.Sc. (Engineering),
Professor, lyudmillaa@mail.ru
Vadim V. Kazakov¹, Postgraduate Student,
kazakov-vadim-k@yandex.ru
Boris V. Karyshev¹, Student,
boris.karyshev@gmail.com

¹ Bauman Moscow State Technical University,
Moscow, 105005, Russian Federation