科学研究費助成事業 研究成果報告書

令和 3 年 6 月 1 8 日現在

機関番号: 62615

研究種目: 基盤研究(B)(一般)

研究期間: 2018~2020

課題番号: 18日03296

研究課題名(和文)Data Mining for Graphs and Networks via Local Intrinsic Dimensional Modeling

研究課題名(英文)Data Mining for Graphs and Networks via Local Intrinsic Dimensional Modeling

研究代表者

Michael E. Houle (Houle, Michael)

国立情報学研究所・大学共同利用機関等の部局等・客員教授

研究者番号:90399270

交付決定額(研究期間全体):(直接経費) 13,100,000円

研究成果の概要(和文):このコンピュータサイエンスプロジェクトの全体的な目的は、特徴アンサンブル、グラフ、ネットワークなどの複雑なデータタイプの類似性アプリケーションにおける次元の呪いの問題に取り組むことであった。 主な目標:(1)組み合わせデータ型を説明するために内在次元の既存の理論を進歩させること。(2)実証的研究によって理論的意味を確認する。(3)ローカル内在次元ににして、新しい進歩を活用して、データベース、データマイニング、マルチメディア、特にグラフ、ディープニューラルネットワーク、フィーチャアンサンブル、その他の複雑なデータタイプのアプリケーション向けのために効率的で効果的なソリューションを開発する。

研究成果の学術的意義や社会的意義 この研究プロジェクトは、機械学習とデータマイニング分野のトップ国際会議(ICML、ICMR、KDD、SDM)において、5つの影響力のある出版物を生み出した。この研究は、実際に非常に大きな影響を及ぼし、3年未満で、2つの論文が、合計約500件近く引用されるまでに達した。

研究成果の概要(英文): The overall aim of this computer science project was to tackle the problem of the curse of dimensionality in similarity applications for complex data types such as feature ensembles, graphs and networks, through the further development of practical techniques that recognize and take advantage of local variations in the intrinsic dimensionality of the data. The main goals were: (1) to advance the existing theory of intrinsic dimensionality to account for combinatorial data types; (2) to confirm the theoretical implications by means of empirical study; (3) to exploit new advances in the theory of local intrinsic dimensionality (LID) to develop more efficient and effective solutions for applications of databases, data mining and multimedia, particularly for graphs, deep neural networks, feature ensembles, and other complex data types.

研究分野: data mining, machine learning, similarity search

キーワード: 高次元空間 極値理論 データマイニング 機械学種 ニューラルネットワーク

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

In the era of Big Data, data volumes have become so enormous and so complex as to preclude processing using traditional applications. For similarity search and retrieval, as well as many other fundamental operations in such areas as data mining, machine learning, multimedia, recommendation systems, and bioinformatics, the efficiency and effectiveness of software implementations depends crucially on the interplay between measures of data similarity and the features (or attributes) by which data objects are represented. For such applications (which we refer here to as *similarity applications*), features are often sought so as to provide the best possible coverage across a range of anticipated queries. However, for any given query, only a relatively small number of features may turn out to be relevant. When the number of features (the data *dimensionality*) is high, the errors introduced into similarity measurements by the many irrelevant feature attributes can completely overwhelm the contributions of the relevant features. As the dimensionality of data increases, the discriminative ability of similarity measures diminishes to the point where techniques that depend on search lose their effectiveness. This phenomenon, of profound importance to computing, is commonly referred to as the *curse of dimensionality*.

Although high-dimensional data sets are generally difficult to process, it is not always the case. Data sets of high dimensionality can often be searched and analyzed relatively easily, provided that the so-called *intrinsic dimensionality* is low. The intrinsic dimensionality of a dataset can loosely be regarded as the number of latent variables needed to describe a data set: this could be the number of 'relevant' or 'non-redundant' features, or the degrees of freedom of a distribution that accurately models the data. Over the past decades, very many formal characterizations of the intrinsic dimensionality of data sets have been proposed within the statistical and machine learning research communities.

Until the beginning of the project, the theory and practice of intrinsic dimensionality had been limited to global approaches, where the intrinsic dimensionality is expressed as a single value for an entire dataset. However, this does not fit well with the user's experience of the data, where interest may be focused on a specific query result or a specific data cluster. Just as data is often modeled as a mixture of underlying distributions, data sets have regions of varying intrinsic dimensionality – and these local variations should be modeled theoretically and exploited in practice.

In 2013 the *local intrinsic dimension* (LID) was introduced so as to model the behavior of distance distributions within the vicinity of points of interest (such as query objects in indexing, or test objects in machine learning). The LID model is comprehensive and general, and practical estimators for LID existed. Since its first appearance, LID had already made an impact in practice, in guiding the design and performance analysis of a wide variety of state-of-the-art solutions for similarity applications.

The LID model is a model of complexity that assesses growth rate in continuous processes, such as of expected neighborhood sizes in similarity applications, or of probability measure in statistical modeling. Despite its successes with continuous data models, LID in its current form cannot be directly applied to discrete data such as graphs, networks, or lists. There are many important application areas where the natural data representations are discrete, and for which there are no satisfactory measures of local query result complexity. For social networks and deep neural network classifiers represented as graphs, or recommender systems and search systems that aggregate ranked lists of data objects or collections of feature sets, direct measures of the complexity of similarity query results are difficult to compute and to apply, due to their reliance on the discrete combinatorial structure of the underlying data representations. For this reason, researchers have been interested in developing continuous generation models for certain classes of structured data, where instances can be thought of as having been constructed according to a particular random generation scheme. Examples of random generation include: a form of densely-connected graph for cluster structure; models of scale-free graphs for social networks; and infinite-neuron layer representations for the theoretical analysis of certain aspects of deep neural network classification.

The main scientific questions to be addressed by this project were as follows. For applications in machine learning, data mining, and other areas in which similarity information is used, how can important structured data types be modeled continuously so as to best support generalization from instances (such as in unsupervised learning from data sets, or supervised learning from training sets)? And can the theory of local intrinsic dimensionality be extended so as to exploit these models in practice?

2. 研究の目的

The overall goal of this project was to tackle the curse of dimensionality in similarity applications for structured data types such as feature ensembles, ranked lists, graphs and networks, through the further development of practical techniques take advantage of local variations in the intrinsic dimensionality of the

data. This was to be done through the extension of the LID theory so as to account for continuous models of structured data and the interactions among combination of features, followed by exploitation of the extended LID model to improve the performance of similarity applications on such data beyond that of current state-of-the-art methods. Within this context, the goals of the project can be summarized as follows:

- To advance the existing theory of intrinsic dimensionality to account for complex data types, in which
 continuous generative models are developed and analyzed for such types as feature ensembles, graphs
 and classification networks;
- 2. To confirm the theoretical implications by means of empirical study of the relationship between local ID and the performance issues that surround similarity applications;
- 3. To exploit new advances in the theory of LID to develop more efficient and effective solutions for applications of databases, data mining and multimedia, particularly for graphs, deep neural networks, feature ensembles, and other complex data types;
- 4. To make these technical innovations available to researchers and practitioners through their integration into practical systems for search and analysis of data, as well as through publication in top-level international journals and conference proceedings.
- 5. To further promote the interdisciplinary international research community focused on the issues of dimensionality and scalability in the context of databases, machine learning, data mining, and multimedia.

3. 研究の方法

The project was a 3-year international collaboration among 6 main participants (the Principal Investigator plus 5 international Research Collaborators), each one a member of a different research institute or university. Collectively, the participants have expertise that spans the areas of computer science that are touched by the curse of dimensionality, including: theory, algorithms, and similarity applications for databases, machine learning and AI (including deep learning), data mining, and multimedia (search and indexing, clustering, outlier detection, feature selection). Within the first year, we planned to expand the theory of LID to account for the interactions among combinations of features, and to show that decompositions of weighted LID can be used to determine low-dimensional local data projections using an approach resembling that of PCA. In the area of graphs and networks in learning contexts, the goal for the first year was to investigate the phenomenon of adversarial attack and deep neural network classification. In later years of the project, we planned to extend the work to follow-on projects in intrinsic dimensionality for complex data applications, with a particular emphasis on deep learning and graph analysis. In all three years of the project, new techniques were to first be applied to simple continuous data types where appropriate, so as to assess the feasibility of extending them to complex data types.

The following lists the expertise that each member brought to the project when it began in 2018 (current affiliations and positions are shown).

- Vis. Prof. Michael E. Houle, National Institute of Informatics (NII), Japan, PI: theoretical computer science, algorithms, data mining, indexing, similarity applications, computational geometry. His main expertise was in the design and analysis of algorithms for similarity search, clustering, outlier detection and feature selection that exploit intrinsic dimensionality.
- Dr. Laurent Amsaleg, CNRS-IRISA Rennes, France, RC: multimedia databases, multidimensional indexing, content-based retrieval, database security and privacy. Dr. Amsaleg had developed practical multimedia indexing systems handling billions of images, represented by millions of features. He had collaborated with Houle on the development of estimators of LID as well as other similarity applications.
- Assoc. Prof. Miloš Radovanović, University of Novi Sad, Serbia, RC: data mining, clustering, experimental validation. Prof. Radovanović is the foremost authority on the hubness phenomenon (implicitly addressed by the LID model), as well as algorithms for data clustering which exploit it.
- Prof. James Bailey, University of Melbourne, Australia, RC: data mining, machine learning, clustering, outlier detection. Prof. Bailey had recently collaborated with Houle, Amsaleg, and Radovanović on an analysis of adversarial attack, and with Houle on measuring functional dependency, both using LID.
- Prof. Arthur Zimek, University of Southern Denmark (USD), Denmark, RC: data mining, clustering, outlier
 detection. Prof. Zimek is a leading authority in these areas, having co-authored several prominent
 surveys. Prof. Zimek had had several joint publications with the PI in the area of data mining. While at
 the University of Munich (LMU) in Germany, Prof. Zimek co-created the well-known ELKI data mining
 application development platform.
- Prof. Vincent Oria, New Jersey Institute of Technology (NJIT), USA, RC: databases, multimedia indexing.
 Prof. Oria had collaborated closely with the PI on applications of ID to adaptive similarity queries and subspace similarity search.

Each of these main team members was responsible for organizing contributions to the project drawn from their own research collaborators and students.

Of the goals laid out at the beginning of the project, we have fully met our objectives with respect to the 3 main research goals. The 4th and 5th goals were only partially met, due to the severe difficulties in implementation and publication during the last year of the project (due to illness, to the lack of availability of students supervised by project members, and on international travel restrictions). In particular, we were not able to fully meet our objectives of integrating our research results into practical systems (due to the lack of availability of students), and in hosting a third international meeting on the topic of LID and its applications (due to the restrictions on travel).

The main discovery of the project is that the local ID model can be used to guide and explain, in a very useful and intuitive way, some of the important characteristics of learning and classification using deep neural networks. The most important outcomes with the greatest impact to date are:

- Publication of a refereed paper at the top-tier International Conference on Learning Representations (ICLR 2018). This paper presented a characterization of corrupted examples in adversarial attack on classification systems, and a practical detection method, based on the LID model central to this project. (Acceptance rate: 2.5%. Number of citations to date, in 3 years: 317.)
- Publication of a refereed paper at the top-tier International Conference on Machine Learning (ICML 2018). This paper demonstrated that the progress of learning in deep neural network (DNN) classifiers is strongly correlated with a drop in LID at the deep feature level, and showed how to use this effect to prevent overtraining and overfitting to data. (World first in automatic detection and avoidance of overtraining during DNN learning. Number of citations to date, in 3 years: 172.)
- Publication of a refereed paper at the top-tier ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2019). This paper presented a technique for generating more realistic synthetic neighbors for the purpose of explainability of learned solutions in deep learning. The improvements to explainability were due to the use of the local intrinsic dimensionality (LID) model central to this project. (Acceptance rate: 14%.)
- Publication of a refereed paper at the top-tier SIAM International Conference on Data Mining (SDM 2019). This paper proposed an estimator for LID that makes use of all pairwise distances within a local neighborhood sample. This use of full similarity neighborhoods allows significantly higher quality of estimation within more tightly-focused neighborhoods. This estimator targets applications in anomaly and outlier detection in data mining, among others. (Acceptance rate: 23%.)
- A refereed publication at a top international conference (SDM 2021) on the use of local intrinsic dimensionality (LID) on the problem of out of-distribution detection in deep learning. LID is used as a feature to characterize the data subspaces formed by data samples and their corresponding reconstruction by autoencoders.
- A refereed publication in early 2021 in the top journal IEEE Transactions on Information Forensics
 and Security, in which the full details of a theoretical analysis is given showing that the vulnerability
 of classification to adversarial perturbation increases as the local intrinsic dimensionality rises.

Other highlights within the project outcomes include:

- 2018: Refereed publications at international conferences: (3 at SISAP 2018) correlation between LID and outlierness, the use of LID in accelerating the performance of data fingerprinting, and an adaptation of LID to model the local growth rate of search neighborhoods within graphs; (1 at WIMS 2018) examination of the effect of reverse neighborhood imbalance in similarity graph construction.
- 2019: 2 refereed publications at international venues: (SISAP 2019 conference) the use of decompositions of LID to determine relevant local data subspaces; (IJAIT journal) an empirical analysis of the NN-descent similarity search method, in terms of local neighborhood imbalance (the hubness phenomenon, related to extreme LID scores). 2 unrefereed publications (arXiv): a full version of the SISAP 2019 paper; and work in progress on the use of an LID-based regularization for improving the quality of GAN-based deep learning.
- A refereed international conference publication (at SISAP 2020) on a theoretical aspect of the LID model, on the use of similarity measures in determining density ratios between two localities in distance distributions. This work lays the first theoretical and practical foundations for density-based outlier detection that properly takes into account the effects of variation in local intrinsic dimension. Preliminary implementations show very competitive performance against the state of the art; however, the work has not yet been published.

The funding for this project was intended to support collaborative research visits among partners from Japan and 5 countries (Australia, Denmark, France, Serbia, and the United States), and conference travel for the reporting of research results. However, as travel has been disrupted due to the COVID-19 pandemic, the research funds for 2020 have gone unused, and (due to difficult circumstances for most of the participants) projects have been delayed. For this reason, most of the projects planned for 2020 will fully resume only after the end of the pandemic. Nevertheless, we are pleased that the project was able to meet its main objectives regardless.

5 . 主な発表論文等

〔雑誌論文〕 計18件(うち査読付論文 16件/うち国際共著 16件/うちオープンアクセス 0件)

〔雑誌論文〕 計18件(うち査読付論文 16件/うち国際共著 16件/うちオープンアクセス 0件)	
1.著者名	4 . 巻
Wang Qizhou, Erfani Sarah M., Leckie Christopher, Houle Michael E.	21
wang Q12100, E11am Sarah w., Leckie om istopher, noure withaer L.	
2 . 論文標題	5 . 発行年
A Dimensionality-Driven Approach for Unsupervised Out-of-distribution Detection	2021年
2 4854-67	6 見知に見後の百
3.雑誌名	6.最初と最後の頁
SIAM International Conference on Data Mining (SDM 2021)	118-126
* * * * * * * * * * * * * * * * * * * *	
掲載論文のDOI(デジタルオブジェクト識別子)	本共の大畑
	査読の有無
10.1137/1.9781611976700.14	有 有
オープンアクセス	国際共著
· · · · · · =· ·	
オープンアクセスではない、又はオープンアクセスが困難	該当する
1.著者名	4 . 巻
	_
Amsaleg Laurent, Bailey James, Barbe Amelie, Erfani Sarah M., Furon Teddy, Houle Michael E.,	16
Radovanovic Milos, Nguyen Xuan Vinh	<u> </u>
2.論文標題	5.発行年
High Intrinsic Dimensionality Facilitates Adversarial Attack: Theoretical Evidence	2021年
mgn intrinsic philensionality racifitates adversalial attack. Inedictical evidence	2021+
3.雑誌名	6.最初と最後の頁
IEEE Transactions on Information Forensics and Security	854-865
TELE Transactions on information foreigness and security	034-003
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1109/TIFS.2020.3023274	有
10.1109/1115.2020.3023274	F F
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
4 *****	1 a 24
1.著者名	4 . 巻
1 . 著者名 Houle Michael E.	4.巻 13
	_
Houle Michael E.	13
Houle Michael E. 2 .論文標題	5 . 発行年
Houle Michael E.	13
Houle Michael E. 2 .論文標題	5 . 発行年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity	13 5.発行年 2020年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名	13 5 . 発行年 2020年 6 . 最初と最後の頁
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity	13 5.発行年 2020年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名	13 5 . 発行年 2020年 6 . 最初と最後の頁
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	13 5 . 発行年 2020年 6 . 最初と最後の頁
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	5 . 発行年 2020年 6 . 最初と最後の頁 248-260
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	5 . 発行年 2020年 6 . 最初と最後の頁 248-260
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有
Houle Michael E. 2. 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3. 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	13 5. 発行年 2020年 6. 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1.著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos,	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19
Houle Michael E. 2.論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3.雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1.著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos,	13 5.発行年 2020年 6.最初と最後の頁 248-260 査読の有無 有 国際共著 該当する
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) #### ### ### ### ### ### ### ### ###	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁
Houle Michael E. 2 .論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 .雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 .著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 .論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 .雑誌名 SIAM International Conference on Data Mining (SDM 2019)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名 SIAM International Conference on Data Mining (SDM 2019) 掲載論文のDOI(デジタルオブジェクト識別子)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189
Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名 SIAM International Conference on Data Mining (SDM 2019)	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) #### ### ### ### ### ### ### ### ###	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名 SIAM International Conference on Data Mining (SDM 2019) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1137/1.9781611975673.21	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189 査読の有無 有
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI (デジタルオプジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名 SIAM International Conference on Data Mining (SDM 2019) 掲載論文のDOI (デジタルオプジェクト識別子) 10.1137/1.9781611975673.21	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189 査読の有無 有
### Houle Michael E. 2 . 論文標題 Local Intrinsic Dimensionality III: Density and Similarity 3 . 雑誌名 International Conference on Similarity Search and Applications (SISAP 2020) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1007/978-3-030-60936-8_19 オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Amsaleg Laurent, Chelly Oussama, Houle Michael E., Kawarabayashi Ken-ichi, Radovanovic Milos, Treeratanajaru Weeris 2 . 論文標題 Intrinsic Dimensionality Estimation within Tight Localities 3 . 雑誌名 SIAM International Conference on Data Mining (SDM 2019) 掲載論文のDOI(デジタルオブジェクト識別子) 10.1137/1.9781611975673.21	13 5 . 発行年 2020年 6 . 最初と最後の頁 248-260 査読の有無 有 国際共著 該当する 4 . 巻 19 5 . 発行年 2019年 6 . 最初と最後の頁 181-189 査読の有無 有

4	A '11'
1.著者名 Jia Yunzhe, Bailey James, Ramamohanarao Kotagiri, Leckie Christopher、Houle Michael E.	4.巻 25
The full life, buffley dumos, rumumonanarao notagiff, Loonto off Totophor, floare infoliaer L.	
2.論文標題	5 . 発行年
Improving the Quality of Explanations with Local Embedding Perturbations	2019年
3.雑誌名	6.最初と最後の頁
ACM SIGKDD International Conference on Knowledge Discovery & Data Mining	875-884
掲載論文のDOI(デジタルオプジェクト識別子)	
10.1145/3292500.3330930	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
	1
1 . 著者名	4.巻
Becker Ruben, Hafnaoui Imane, Houle Michael E., Li Pan, Zimek Arthur	12
2.論文標題	5.発行年
Z . 晶光标题 Subspace Determination Through Local Intrinsic Dimensional Decomposition	2019年
outspace betermination intough cocal intrinsic bimensional becomposition	2019-
3 . 雑誌名	6.最初と最後の頁
International Conference on Similarity Search and Applications (SISAP 2019)	281-289
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1007/978-3-030-32047-8_25	有
オープンアクセス	国際共著
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国际共省 該当する
カープンプラと人にはない、人はカープンプラと人が四無	成当する
1.著者名	4.巻
Bratic Brankica, Houle Michael E., Kurbalija Vladimir, Oria Vincent, Radovanovic Milos	28
•	
2.論文標題	5 . 発行年
The Influence of Hubness on NN-Descent	2019年
2 hhttp://	C 目初1.目後の五
3.雑誌名	6.最初と最後の頁
International Journal on Artificial Intelligence Tools	1960002-1960002
掲載論文のDOI(デジタルオプジェクト識別子)	査読の有無
10.1142/\$0218213019600029	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
4 英北久	
1. 著者名	4.巻 abs/1907.06771
Becker Ruben, Hafnaoui Imane, Houle Michael E., Li Pan, Zimek Arthur	aus/130/.00//1
2 . 論文標題	5.発行年
Subspace Determination through Local Intrinsic Dimensional Decomposition: Theory and	2019年
Experimentation	
3.雑誌名	6.最初と最後の頁
arXiv CoRR	1-17
掲載論文のDOI(デジタルオブジェクト識別子)	本性の左無
何製論スリルリしてングルタノンエンで蔵別士)	査読の有無
なし	無
なし	
	国際共著

1.著者名	4 . 巻
Barua Sukarna, Ma Xingjun, Erfani Sarah M., Houle Michael E., Bailey James	abs/1905.00643
2.論文標題	5.発行年
Quality Evaluation of GANs Using Cross Local Intrinsic Dimensionality	2019年
3.雑誌名	6.最初と最後の頁
arXiv Corr	1-31
and continued to the co	
掲載論文のDOI(デジタルオブジェクト識別子)	<u></u> 査読の有無
なし	無
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	-
1 . 著者名	4 . 巻
Yunzhe Jia, James Bailey, Kotagiri Ramamohanarao, Christopher Leckie, Michael E. Houle	4 · 문 25
2 . 論文標題	5.発行年
Improving the Quality of Explanations with Local Embedding Perturbations	2019年
3 . 雑誌名	6.最初と最後の頁
25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2019)	印刷中
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
なし	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
a Table	1 4 4 4
1.著者名 Laurent Amsaleg, Oussama Chelly, Michael E. Houle, Ken-ichi Kawarabayashi, Milos Radovanovic, Weeris Treeratanajaru	4.巻 19
2 . 論文標題	5.発行年
Intrinsic Dimensionality Estimation within Tight Localities	2019年
3.雑誌名	6.最初と最後の頁
19th SIAM International Conference on Data Mining (SDM 2019)	181 ~ 189
担要なみのDOL / ごごクリナイジュカト 幼叫フト	本共の左便
掲載論文のDOI(デジタルオブジェクト識別子) 10.1137/1.9781611975673.21	査読の有無 有
	C Dim 11 +++
オープンアクセス	国際共者
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
オープンアクセスではない、又はオープンアクセスが困難	該当する
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi	
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett	該当する 4 . 巻 32
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett	該当する 4 . 巻
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett 2 . 論文標題 Extreme-value-theoretic estimation of local intrinsic dimensionality	該当する 4 . 巻 32 5 . 発行年 2018年
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett 2 . 論文標題 Extreme-value-theoretic estimation of local intrinsic dimensionality	該当する 4 . 巻 32 5 . 発行年
1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett 2 . 論文標題 Extreme-value-theoretic estimation of local intrinsic dimensionality 3 . 雑誌名 Data Mining and Knowledge Discovery	該当する 4 . 巻 32 5 . 発行年 2018年 6 . 最初と最後の頁 1768~1805
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett 2 . 論文標題 Extreme-value-theoretic estimation of local intrinsic dimensionality 3 . 雑誌名 Data Mining and Knowledge Discovery	該当する 4 . 巻 32 5 . 発行年 2018年 6 . 最初と最後の頁
オープンアクセスではない、又はオープンアクセスが困難 1 . 著者名 Laurent Amsaleg, Oussama Chelly, Teddy Furon, Stephane Girard, Michael E. Houle, Ken-ichi Kawarabayashi, Michael Nett 2 . 論文標題 Extreme-value-theoretic estimation of local intrinsic dimensionality 3 . 雑誌名 Data Mining and Knowledge Discovery	該当する 4 . 巻 32 5 . 発行年 2018年 6 . 最初と最後の頁 1768~1805

1.著者名	4 . 巻
Lorenzo von Ritter, Michael E. Houle, Stephan Guennemann	11
2 - 全人中 4面 日百	r 28/=/=
2.論文標題	5 . 発行年
Intrinsic Degree: An Estimator of the Local Growth Rate in Graphs	2018年
3.雑誌名	6.最初と最後の頁
11th International Conference on Similarity Search and Applications (SISAP 2018)	195 ~ 208
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1007/978-3-030-02224-2_15	有
10.1007/970-3-030-02224-2_13	i i i i i i i i i i i i i i i i i i i
4	
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
	<u>. </u>
1 . 著者名	4 . 巻
	_
Michael E. Houle, Erich Schubert, Arthur Zimek	11
2.論文標題	5.発行年
	2018年
On the Correlation Between Local Intrinsic Dimensionality and Outlierness	2010 1
- 404 5	
3.雑誌名	6.最初と最後の頁
11th International Conference on Similarity Search and Applications (SISAP 2018)	177 ~ 191
· · · · · · · · · · · · · · · · · · ·	
	本共の大畑
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1007/978-3-030-02224-2_14	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
オーノンテクと人とはない、又はオーノンテクと人が四無	政当りる
1.著者名	4 . 巻
Michael E. Houle, Vincent Oria, Kurt R. Rohloff, Arwa M. Wali	11
2 . 論文標題	5.発行年
LID-Fingerprint: A Local Intrinsic Dimensionality-Based Fingerprinting Method	2018年
3.雑誌名	6.最初と最後の頁
11th International Conference on Similarity Search and Applications (SISAP 2018)	134 ~ 147
Titli international conference on Similarity Search and Applications (SISAF 2016)	134 ** 147
掲載論文のDOI(デジタルオブジェクト識別子)	査読の有無
10.1007/978-3-030-02224-2_11	有
13.135.73.73 3 300 SEEE 2_11	7
+ -f\75.4	同咖 井茶
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
	•
1 . 著者名	4 . 巻
—	_
Xingjun Ma, Yisen Wang, Michael E. Houle, Shuo Zhou, Sarah M. Erfani, Shu-Tao Xia, Sudanthi	35
Wijewickrema, James Bailey	
2 . 論文標題	5 . 発行年
Dimensionality-Driven Learning with Noisy Labels	2018年
Dimensionality bitter Educating with here's Education	2010—
2 145+47	6 見切に見後の下
3.雑誌名	6.最初と最後の頁
35th International Conference on Machine Learning (ICML 2018)	3361 ~ 3370
掲載論文のDOI(デジタルオブジェクト識別子)	本誌の右無
	査読の有無
なし	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
カーファティにかてはない、 入はカーフファブ ピグが凶 無	マスログ

1 . 著者名 Brankica Bratic, Michael E. Houle, Vladimir Kurbalija, Vincent Oria, Milos; Radovanovic	4. 巻 8
·	
2.論文標題	5 . 発行年
NN-Descent on High-Dimensional Data	2018年
3.雑誌名	6.最初と最後の頁
8th International Conference on Web Intelligence, Mining and Semantics (WIMS 2018)	20:1 ~ 20:8
掲載論文のDOI (デジタルオブジェクト識別子)	査読の有無
10.1145/3227609.3227643	有
オープンアクセス	国際共著
オープンアクセスではない、又はオープンアクセスが困難	該当する
1.著者名	4 . 巻
Xingjun Ma, Bo Li, Yisen Wang, Sarah M. Erfani, Sudanthi N. R. Wijewickrema, Grant Schoenebeck, Dawn Song, Michael E. Houle, James Bailey	6
2.論文標題	5 . 発行年
Characterizing Adversarial Subspaces Using Local Intrinsic Dimensionality	2018年
3.雑誌名	6.最初と最後の頁
6th International Conference on Learning Representations (ICLR 2018), CoRR abs/1801.02613	1 ~ 15
掲載論文のDOI(デジタルオプジェクト識別子)	査読の有無
なし	有
オープンアクセス	国際共著

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

6.研究組織

о.	. 听九組織		
	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

該当する

オープンアクセスではない、又はオープンアクセスが困難

7.科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
オーストラリア	University of Melbourne			
フランス	CNRS / IRISA Rennes	INRIA / IRISA Rennes		
米国	New Jersey Institute of Technology			
デンマーク	University of Southern Denmark			
セルビア	University of Novi Sad			