Energy-aware paradigm for designing compact data structures

Susana Ladra (University of A Coruña, Spain)

"P.O. FEDER: Unha maneira de facer Europa"

Brief presentation – Susana Ladra (University of A Coruña)

Researcher at the Centre for ICT Research (CITIC)

Data Compression ; Algorithms and Data Structures

Bioinformatics ; Data Mining ; Energy Efficiency

Coordinator of BIRDS project (Bioinformatics and Information Retrieval Data Structures Analysis and Design, MSCA RISE, GA 690941, 2016-2019)

citic

UNIVERSIDADE DA CORUÑA

Energy-aware paradigm for designing compact data structures

Introduction: Compact Data Structures

Data Structures:

Pre-process input data so as to answer (long) series of retrieval or update operations

Want to minimize:

1. Query time
2. Space usage of data structure
3. Time of pre-processing
4. Space for pre-processing

Slide from: https://materials.dagstuhl.de/files/18/18281/18281.RajeevRaman1.Slides.pdf

Energy-aware paradigm for designing compact data structures

Introduction: Compact Data Structures

Data Structures:

Answering queries on data requires an index in addition to the data

Index may be larger than the data

Space usage = “space for data” + “space for index”

redundancy

Slide from: https://materials.dagstuhl.de/files/18/18281/18281.RajeevRaman1.Slides.pdf

Energy-aware paradigm for designing compact data structures

Introduction: Compact Data Structures

Compact Data Structures:

Space usage = “space for data” + “space for index”

low-order term

and perform operations **directly** on it

→for static data structures, often get O(1) time operations

Slide from: https://materials.dagstuhl.de/files/18/18281/18281.RajeevRaman1.Slides.pdf

Energy-aware paradigm for designing compact data structures

Introduction: Compact Data Structures

Examples of compact Data Structures:

Trees: 2 bits per node, allowing navigation (child, parent, subtree…)

Sequences: $n \log \sigma + o(n \log \sigma)$ bits, allowing rank/select/access

Texts: $O(H_k(T)$ + lower order), allowing counting in $O(|P| \log \sigma)$ time

→ Full-text index occupying 20-30% of the text (**self-indexing**)

Web graphs: 1-2 bits per edge

Slide from: https://materials.dagstuhl.de/files/18/18281/18281.RajeevRaman1.Slides.pdf

Energy-aware paradigm for designing compact data structures

Motivation of the project

Big Data / IoT → compact data structures

Compact data structures → space/time trade-off

What about energy?

Energy-aware paradigm for designing compact data structures

Current status

Green Computing → traditionally focuses on hardware

Green Software → gaining importance

Software Quality, Software Design, Software Requirements…

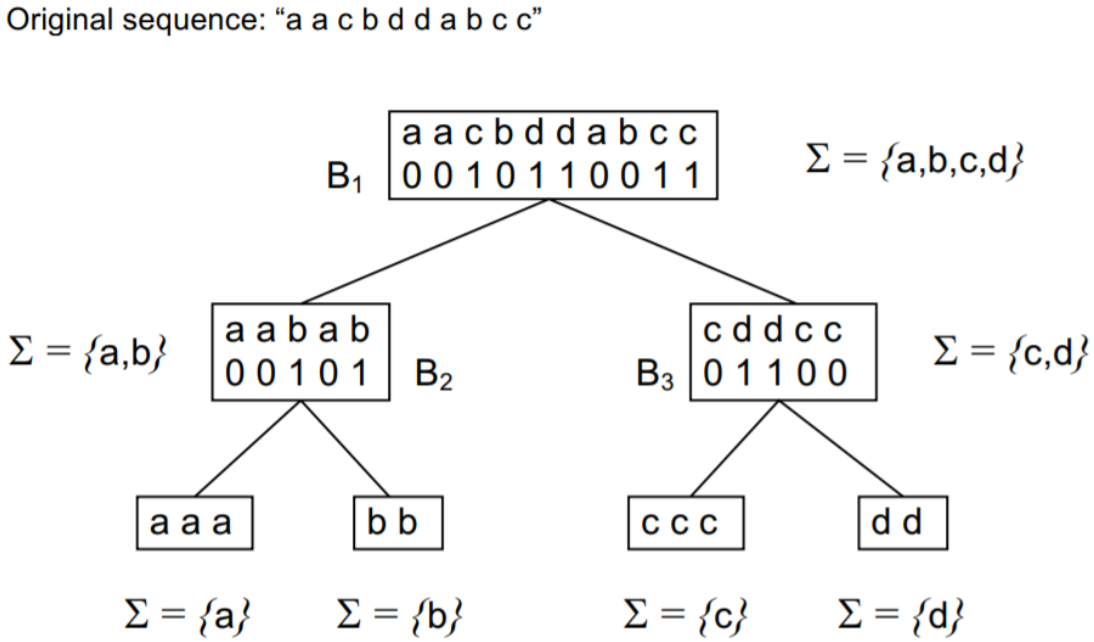
Insufficiently studied from the algorithmic point of view

Not studied at all for compact data structures

Ferragina (2010) made some reflections on the state of the art

Energy-aware paradigm for designing compact data structures

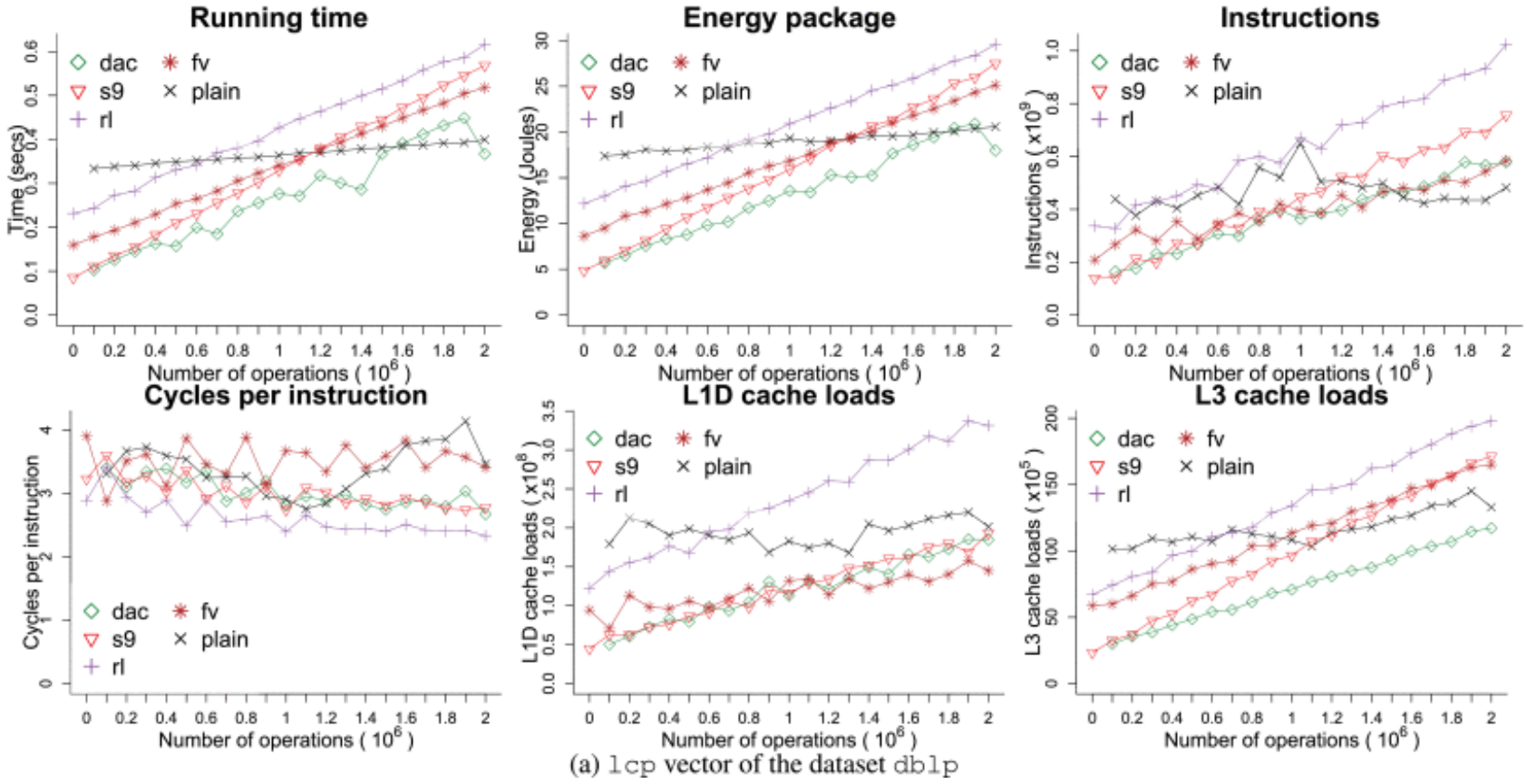
Energy complexity?



Complex memory patterns (not energy-efficient?)

Energy-aware paradigm for designing compact data structures

Preliminary studies



(a) `lcp` vector of the dataset `dblp`

https://doi.org/10.1109/ACCESS.2019.2949655

Susana Ladra (University of A Coruña, Spain)

citic

UNIVERSIDADE DA CORUÑA

Energy-aware paradigm for designing compact data structures

Challenges

→ Better characterize all the factors that impact on the energy consumption

- How to accurately measure the energy consumption?

→ Energy labelling for algorithms and data structures?

Susana Ladra

susana.ladra@udc.es

Universidade da Coruña Campus Universitario de Elviña s/n
15071 A Coruña
T. +34 881 01 5501
www.citic-research.org

"P.O. FEDER: Unha maneira de facer Europa"