A Tightly-coupled XML Encoder-decoder For 3D Data Transaction

Commission III – Spatial Information Management

Siew Chengxi Bernad
Khairul Hafiz Sharkawi

3D GIS Research Lab, Faculty of Geoinformation and Real Estate,
Universiti Teknologi Malaysia

Outline

- Introduction
- The Implementations
- Encoding Results
- Javascript Decoder
- Conclusions
- Acknowledgements

Introduction

- 3D spatial data sharing are getting popular in 3D SDI especially in urban modelling
- Large datasets file size is expected (>=LoD3 ~1Megabytes to 25MB each building)
- 3D data sharing use cases:
  1. Only visualizations?
  2. Visualizations + event-based semantics retrieval?
  3. Semantics + Geometries both presence for analysis

Introduction (cont.)

- Scenario 3 is not practical by sending CityGML data to clients
- Existing compression techniques does not solve the problem well:
  - Non-query-able
  - Non-streaming
  - General usages
  - No partial decompression

Introduction (cont.)

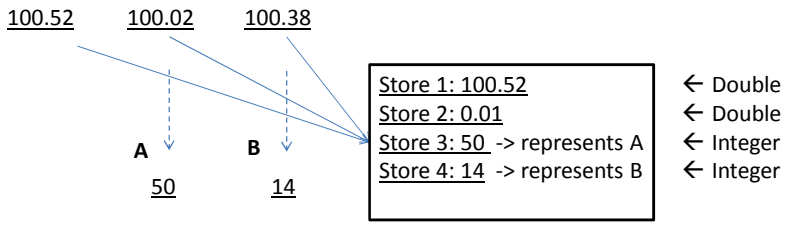
- Encoder-decoder is innovated for this purpose
- This coupling layer allows:
  - Query-able on compressed document
  - Hence allow partial decompression
  - Schema-aware, adapt to large CityGML geometries dataset
- An encoding framework within or with web services for spatial data transfer improves efficiencies

The Implementation

- Encoding based on 7 main components (elements, attributes, attribute values, values, geometries, URI)
- Geometries are encoded based on scaled integer.
- All components are stored in dictionary in uniform 16-bit symbols
- The encoder adapts non-progressive, lossless point, streaming and random access behavior

Scaled Integer

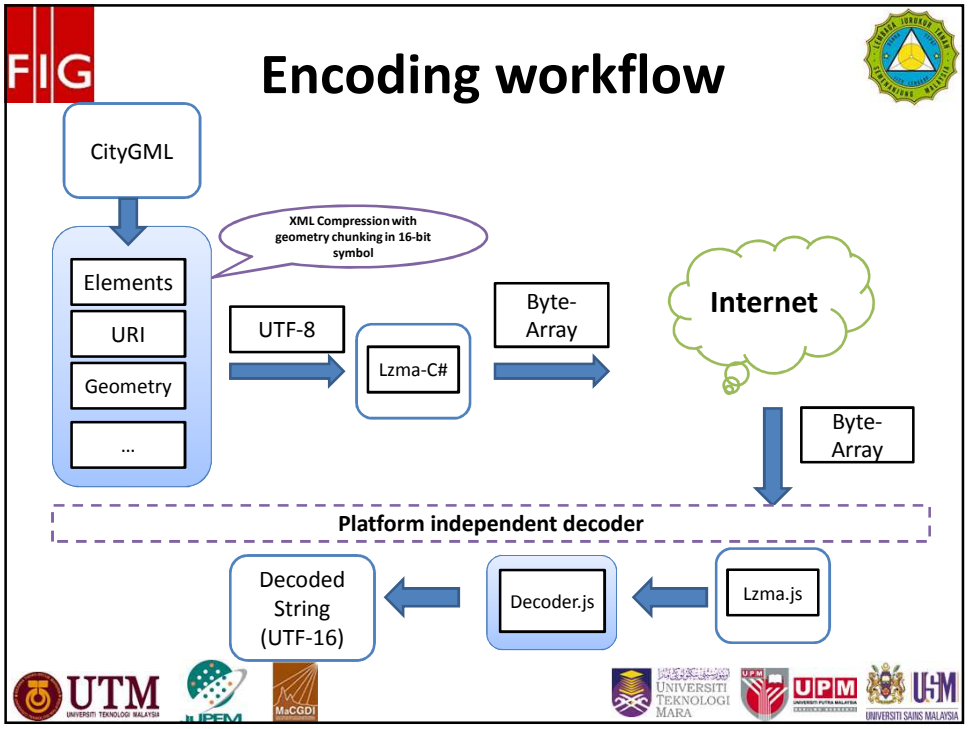
- Scaled Integer (To store geometry information with smaller size)

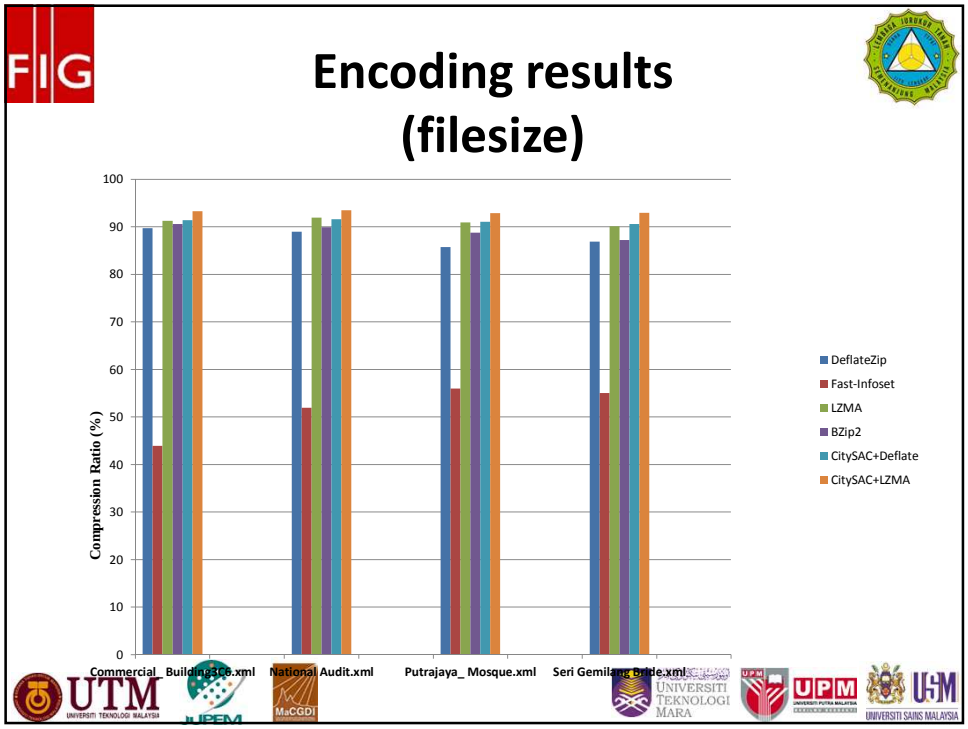


Note: Example by using scaled integer – 100.02 can be stored in integer (4 bytes) instead of double (8 bytes)

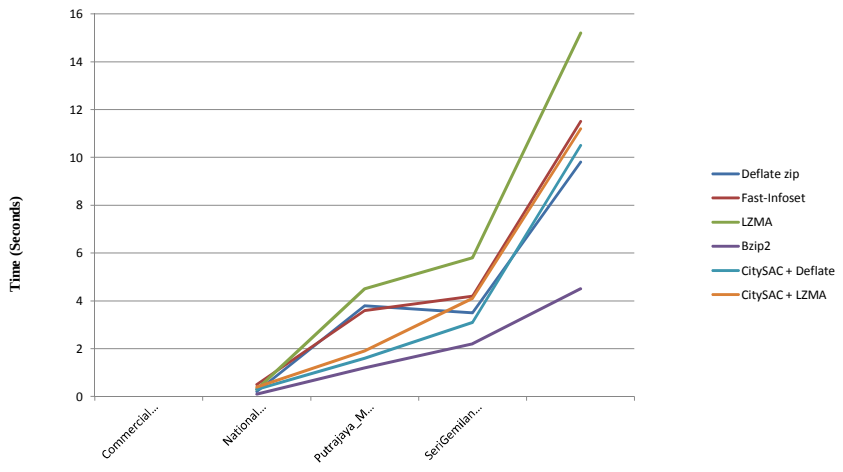
Key components

	Components	format	size
1.	Bit-Representation	u_char	1 byte
2.	Close	u_char	1 byte
3.	Tag	u_char	1 byte
4.	Attribute	u_char	1 byte
5.	Attribute Value	u_char	1 byte
6.	Value	u_char	1 byte
7.	URI	u_char	1 byte
8.	X	int	4 bytes
	Y	int	4 bytes
	Z	int	4 bytes

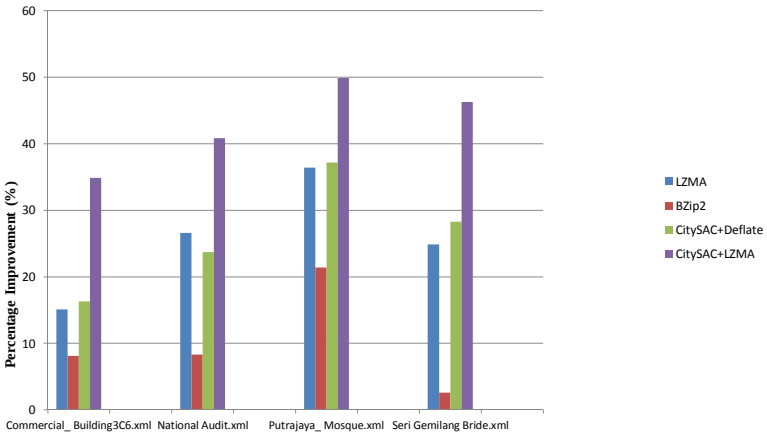

Encoding workflow
CityGML
XML Compression with geometry chunking in 16-bit symbol
Elements
URI
Geometry
...
UTF-8
Lzma-C#
Byte-Array
Internet
Byte-Array
Platform independent decoder
Decoded String (UTF-16)
Decoder.js
Lzma.js


Encoding results (filesize)
Compression Ratio (%)
100
90
80
70
60
50
40
30
20
10
0
Commercial_ Building3C6.xml
National Audit.xml
Putrajaya_ Mosque.xml
Seri Gemilang Bridge.xml
DeflateZip
Fast-Infoset
LZMA
BZip2
CitySAC+Deflate
CitySAC+LZMA

Time taken (seconds)



Improvements (using deflate as benchmark)



Comparisons

Techniques\Characteristics	Query-able	Partial decompression
Deflate alone	No	No
Bzip2 alone	No	No
LZMA alone	No	No
Fast-Infoset	Yes	Yes
CitySAC	Yes	Yes

Encoding environment

Processor	**Intel Core i7-3610M 2.3Ghz**
Memory	6GB DDR3
Hard Disk	1 TB HDD with 34.6GB free space on C:
Operating System	Windows 7 64-bit Professional on DOT.NET Framework 4.5

The Javascript Decoder

- HTML5 and Javascript is becoming popular.
- Javascript provides platform independent solution via interpreter eg. NodeJS and Browser
- Simple binary transformation and decompression could be done.
- Reading binary file is available in WebGL (little-endian and BigEndian)

The Javascript Decoder

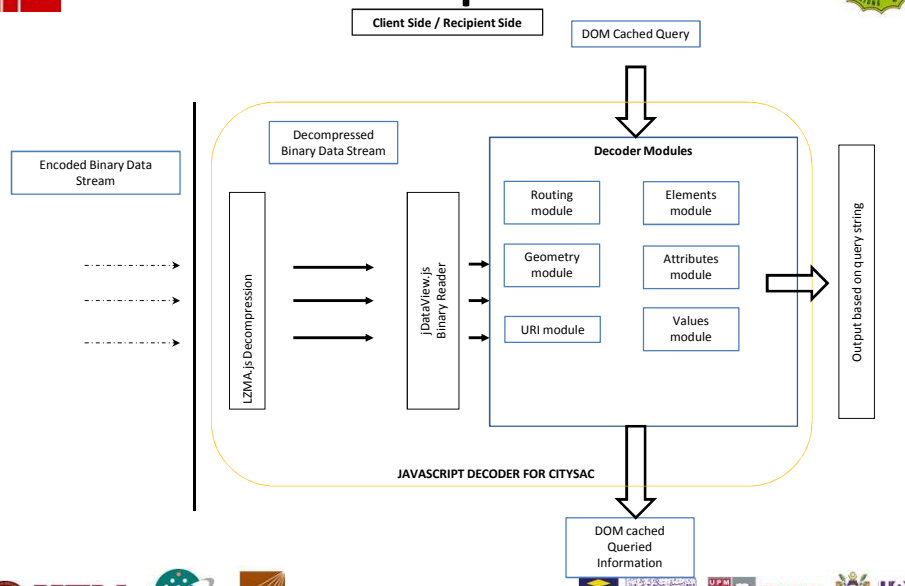
- Query-able modules

Or

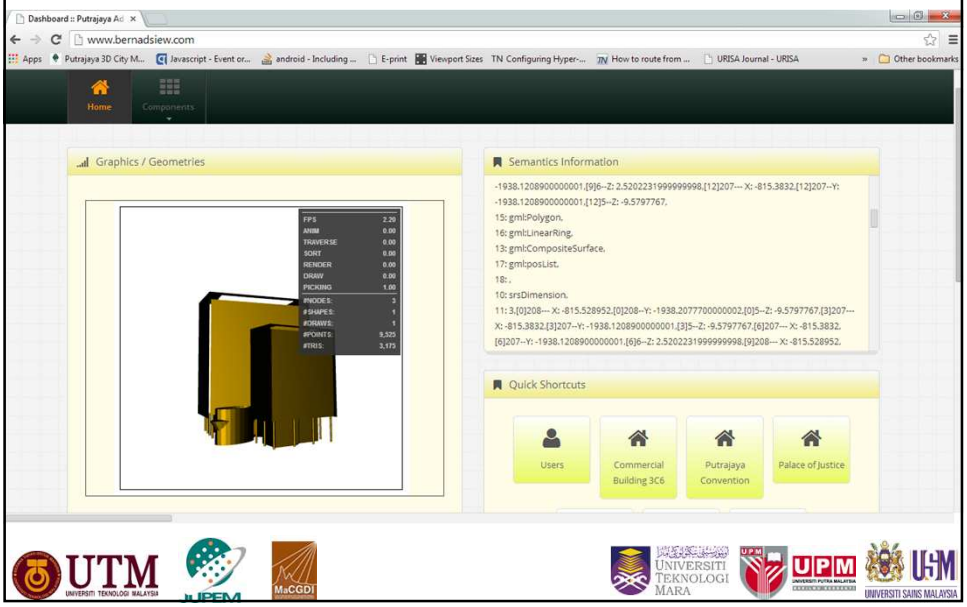
- Full decompression to CityGML

  Compressed -> Partial decompressed (binary data) -> CityGML

The Javascript Decoder



Application snapshots



Conclusions

- Schema-aware encoder able to encode CityGML into smaller size (35% - 50 %) compare to Deflate (WinZIP)
- Schema-aware encoder able to encode CityGML into smaller size (15% - 30 %) and faster compare to LZMA (7-zip)
- Javascript decoder able to provide code-on-demand solution
- Binary representation standard for GML tags?
- XML compression is preferred for large data transaction as partial decompression and query-able compressed document benefits are gained

Acknowledgment

- Financial assistance from Malaysian Peninsular Land Surveyors Board (LJT)
- MyPhD scholarship program from Ministry of Education

Thank you for your attention!

3D GIS Research Lab, FGRE, UTM

Email address:

cbsiew2@live.utm.my

or

bernad@bernadsiew.name