

XML Analysis Template

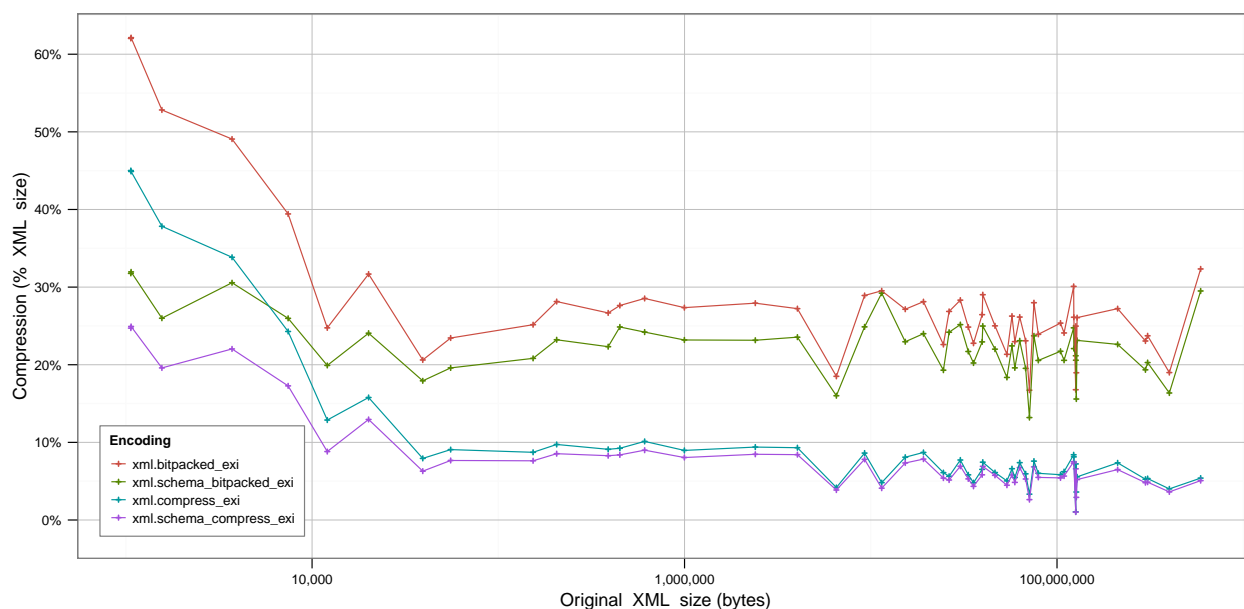
15 January, 2015

Results for Digital Forensics XML (DFXML) Use Case

EXI Exploratory

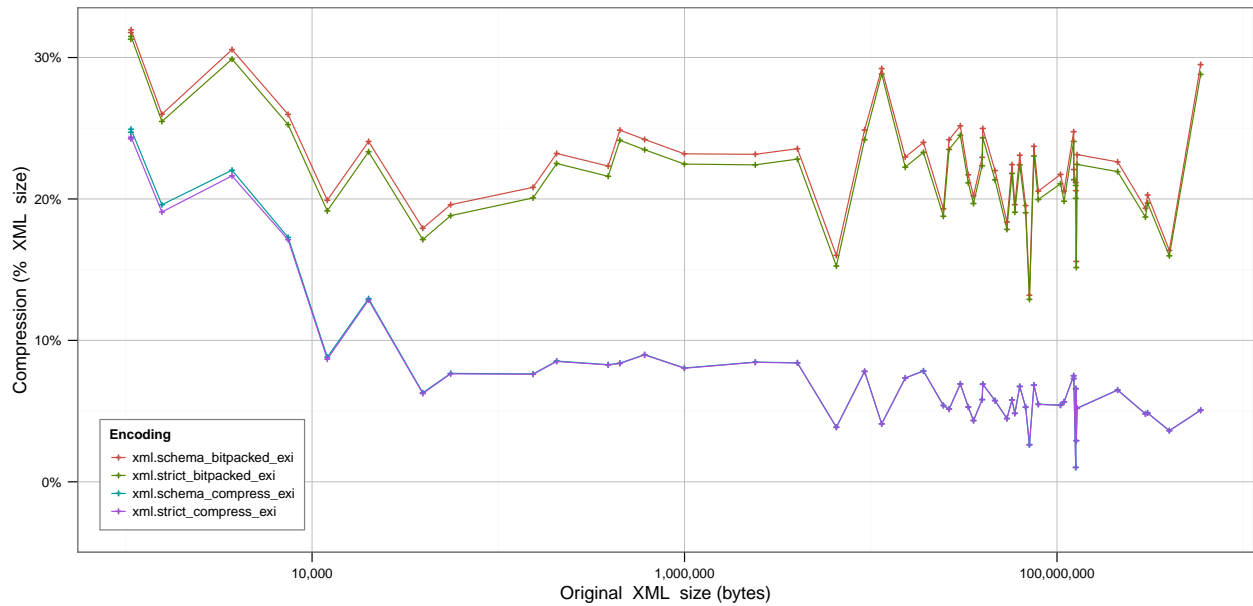
A. How do the primary EXI modes compare for schemaless & schema-informed encodings?

```
## [1] "Series:  xml.bitpacked_exi, xml.schema_bitpacked_exi, xml.compress_exi, xml.schema_compress_exi"
## [1] "Baseline:  xml"
## xml.bitpacked_exi xml.schema_bitpacked_exi xml.compress_exi
## Min.   :0.1672    Min.   :0.1319        Min.   :0.01051
## 1st Qu.:0.2359    1st Qu.:0.2025        1st Qu.:0.05597
## Median :0.2627    Median :0.2262        Median :0.07387
## Mean   :0.2802    Mean   :0.2253        Mean   :0.10029
## 3rd Qu.:0.2823    3rd Qu.:0.2420        3rd Qu.:0.09087
## Max.   :0.6214    Max.   :0.3196        Max.   :0.45019
## xml.schema_compress_exi
## Min.   :0.01020
## 1st Qu.:0.05180
## Median :0.06593
## Mean   :0.07769
## 3rd Qu.:0.08165
## Max.   :0.24930
```



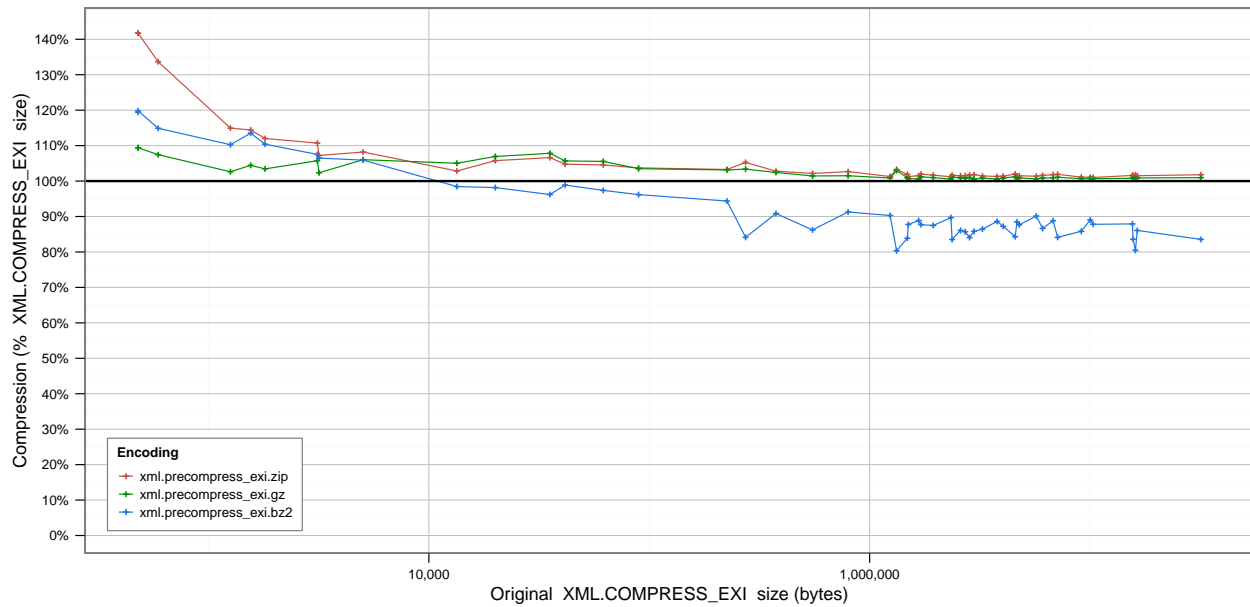
B. Does the 'strict' option significantly improve compaction for schema-informed encodings?

```
## [1] "Series:  xml.schema_bitpacked_ex1, xml.strict_bitpacked_ex1, xml.schema_compress_ex1, xml.strict_compress_ex1"
## [1] "Baseline:  xml"
## xml.schema_bitpacked_ex1 xml.strict_bitpacked_ex1 xml.schema_compress_ex1
## Min.      :0.1319          Min.      :0.1290          Min.      :0.01020
## 1st Qu.:0.2025          1st Qu.:0.1970          1st Qu.:0.05180
## Median :0.2262          Median :0.2193          Median :0.06593
## Mean     :0.2253          Mean     :0.2190          Mean     :0.07769
## 3rd Qu.:0.2420          3rd Qu.:0.2349          3rd Qu.:0.08165
## Max.     :0.3196          Max.     :0.3149          Max.     :0.24930
## xml.strict_compress_ex1
## Min.      :0.01017
## 1st Qu.:0.05160
## Median :0.06580
## Mean     :0.07709
## 3rd Qu.:0.08145
## Max.     :0.24367
```



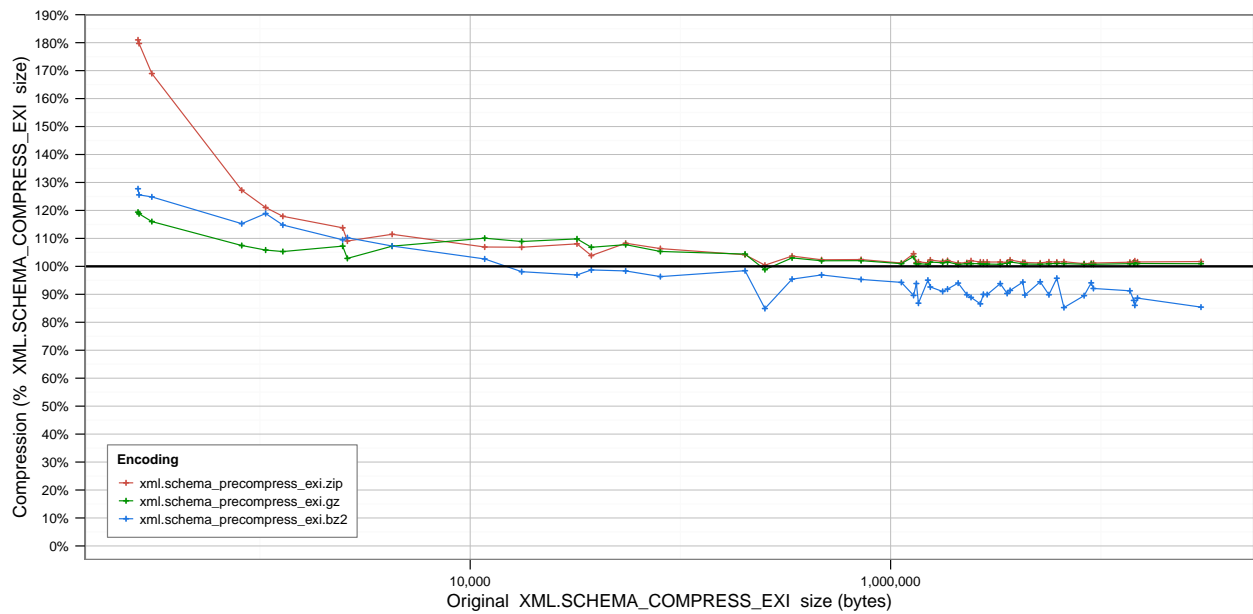
C. Do any of the tested compression algorithms perform better on a schemaless, precompress EXI document than the standard DEFLATE?

```
## [1] "Series:  xml.precompress_exi.zip, xml.precompress_exi.gz, xml.precompress_exi.bz2"
## [1] "Baseline:  xml.compress_exi"
##  xml.precompress_exi.zip  xml.precompress_exi.gz  xml.precompress_exi.bz2
##  Min.    :1.010          Min.    :1.005          Min.    :0.8033
##  1st Qu.:1.015          1st Qu.:1.008          1st Qu.:0.8595
##  Median :1.018          Median :1.010          Median :0.8848
##  Mean   :1.055          Mean   :1.025          Mean   :0.9251
##  3rd Qu.:1.046          3rd Qu.:1.034          3rd Qu.:0.9678
##  Max.   :1.418          Max.   :1.094          Max.   :1.1983
```



D. Do any of the tested compression algorithms perform better on a schema-informed, precompress EXI document than the standard DEFLATE?

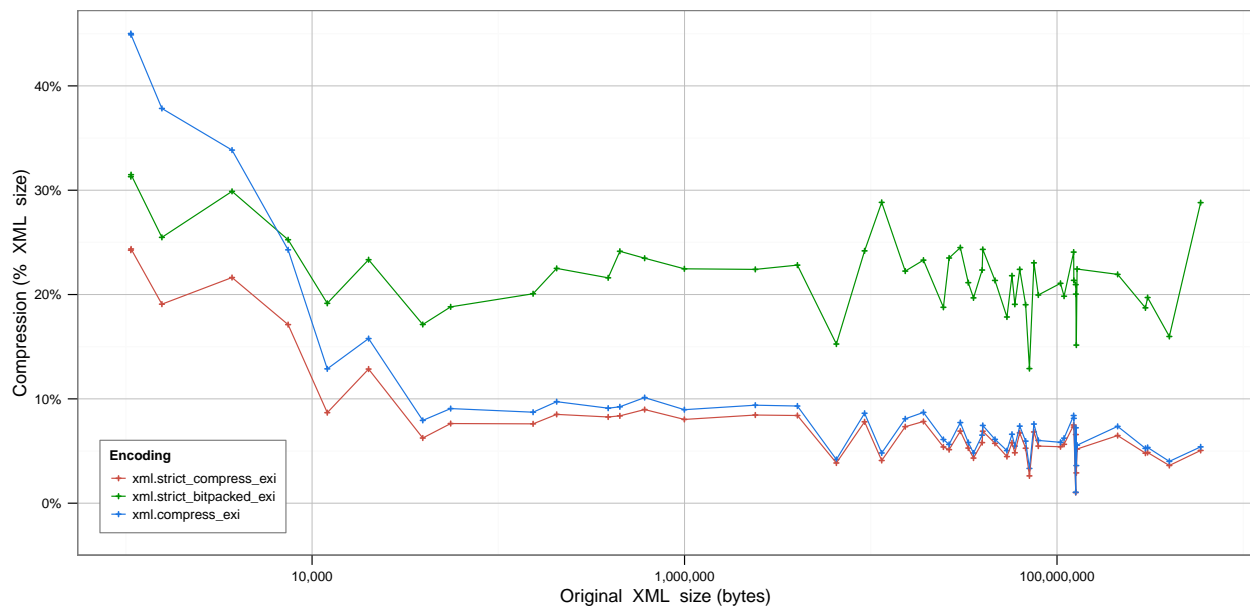
```
## [1] "Series:  xml.schema_precompress_exi.zip, xml.schema_precompress_exi.gz, xml.schema_precompress_exi.bz2"
## [1] "Baseline:  xml.schema_compress_exi"
##  xml.schema_precompress_exi.zip  xml.schema_precompress_exi.gz
##  Min.      :1.004                  Min.      :0.9882
##  1st Qu.:1.015                    1st Qu.:1.0093
##  Median :1.017                    Median :1.0116
##  Mean    :1.085                    Mean    :1.0357
##  3rd Qu.:1.066                    3rd Qu.:1.0531
##  Max.    :1.810                    Max.    :1.1939
##  xml.schema_precompress_exi.bz2
##  Min.      :0.8491
##  1st Qu.:0.8985
##  Median :0.9402
##  Mean    :0.9667
##  3rd Qu.:0.9823
##  Max.    :1.2776
```



Binary-comparisons

E. Which EXI encoding is the most compact?

```
## [1] "Series:  xml.strict_compress_exi, xml.strict_bitpacked_exi, xml.compress_exi"
## [1] "Baseline:  xml"
## xml.strict_compress_exi xml.strict_bitpacked_exi xml.compress_exi
## Min.      :0.01017      Min.      :0.1290      Min.      :0.01051
## 1st Qu.:0.05160      1st Qu.:0.1970      1st Qu.:0.05597
## Median :0.06580      Median :0.2193      Median :0.07387
## Mean    :0.07709      Mean    :0.2190      Mean    :0.10029
## 3rd Qu.:0.08145      3rd Qu.:0.2349      3rd Qu.:0.09087
## Max.    :0.24367      Max.    :0.3149      Max.    :0.45019
```



F. For a network already using gzip, do any of the EXI encodings offer improvements?

```
## [1] "Series:  xml.strict_compress_exi, xml.strict_bitpacked_exi, xml.compress_exi"
## [1] "Baseline:  xml.gz"
##  xml.strict_compress_exi xml.strict_bitpacked_exi xml.compress_exi
##  Min.      :0.05878      Min.      :0.5915      Min.      :0.06076
##  1st Qu.:0.50601      1st Qu.:1.7494      1st Qu.:0.61169
##  Median :0.59030      Median :1.9474      Median :0.66058
##  Mean    :0.56827      Mean    :1.8637      Mean    :0.67228
##  3rd Qu.:0.65522      3rd Qu.:2.1301      3rd Qu.:0.78448
##  Max.    :0.72478      Max.    :3.4527      Max.    :0.95303
```

