Security Considerations Regarding Compression Dictionaries
draft-kucherawy-httpbis-dict-sec-00

Abstract

Data compression algorithms benefit from blocks of tuning data called "dictionaries". These can greatly improve data compression speed and/or ratios, but their use and application has numerous potential security issues of concern to the communities using them. This document enumerates security issues known about compression dictionaries at the time of publication so that future proposals for use of dictionaries can benefit from this collected material.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 9, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction ...................................................... 3
2. Definitions ..................................................... 3
3. Dictionary Security Concerns ................................. 3
4. IANA Considerations ........................................... 4
5. Security Considerations ....................................... 4
6. References ...................................................... 5
   6.1. Normative References ................................. 5
   6.2. Informative References ................................. 5
Appendix A. Acknowledgements ............................... 5
Appendix B. Prior Art ........................................... 5
1. Introduction

Brotli [RFC7932] and Zstandard [RFC8478] are examples of two modern data compression algorithms. While useful in their basic forms, they can be made far more effective with specific types of payloads when used with an object called a "dictionary". A dictionary is a map that can be applied during compression or uncompression that provides an advantage when operating against specific types of content. One might, for example, develop a dictionary that makes the compression algorithm more effective when applied to specific types of audio data.

As dictionaries are being developed, some issues have come to light that indicate ways that use of dictionaries might introduce destructive side effects to the environment in which their use is applied. This document is a collection of those topics, which can be consulted as work on dictionaries progresses; later, as RFCs are published advancing dictionaries, the content of this document could be used as a checklist to ensure that either the algorithms or their specification documents have been appropriately evaluated against these concerns.

2. Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Dictionary Security Concerns

These subsections each describe an issue that has been raised with respect to use of dictionaries as input to compression and uncompression. Where possible and known, acceptable mitigations are described.

[TODO: This is a bullet list for now, but each bullet item will gradually be converted into a subsection containing relevant discussion.]

- Attacks that use dictionary-based compression to recover content in the response.

- Attacks that use dictionary-based compression to recover content in the dictionary.
4. IANA Considerations

This document includes no actions for IANA.

[RFC Editor: Please remove this section before publication.]

5. Security Considerations

This document enumerates known security considerations about a space that is under development. The list of issues discussed above may
not be exhaustive, but it is hopefully complete enough to ensure quality work is produced as a result.

6. References

6.1. Normative References


6.2. Informative References


Appendix A. Acknowledgements

The author wishes to acknowledge the following for their review and constructive criticism of this update: TBD

Appendix B. Prior Art

Some prior art worth considering:

- draft-lee-sdch-spec, which was implemented in Chrome but then withdrawn
- draft-vkrasnov-h2-compression-dictionaries
- draft-vandevenne-shared-brotli-format
- HTTPBIS discussion during IETF 97
- Brotli "fetch spec" proposal: https://fetch.spec.whatwg.org/
o various HTTPBIS mailing list threads about dictionaries

Author’s Address

Murray S. Kucherawy
Facebook, Inc.
1 Hacker Way
Menlo Park, CA 94025
US

EMail: msk@fb.com