

RFCをデータで眺めてみた

~References引用グラフで「読む候補」を探す~

BBSakura Networks 株式会社 外山 結月

自己紹介

- 外山 結月（とやま ゆづき）
 - 所属：BBSakura Networks株式会社 / BBIX株式会社 ITシステム室
- 経歴
 - 2024年4月～現在：BBSakura Networks株式会社 / BBIX株式会社
 - ソフトウェアエンジニア → ネットワーク初心者（RFC初心者）
 - 社内の契約管理システムのバックエンド開発
 - オンプレミスのインフラ運用
- JANOG参加歴
 - JANOG 56, 57, 58→3回目！

システム開発でこんな経験、ありませんか？

「MTAサーバたってるから、SMTPでそこに投げれば外部にメール送れるよ」

→SMTPとは？？MTAとは？？

「SNMPでMIB指定してVLANの情報取得して差分検出したい」

→SNMPとは？？MIBとは？？

Google検索・LLMに質問・本の一節を読み、解決

本当はRFCとか読んでみるべきなんだろうな...

RFC読んでみたいけれど...

- なにから読めば良いのかわからない
- どう読めば良いのかわからない

→ References 引用グラフ + PageRank で解消していく！

方法（引用グラフ）

各RFCのReferences節から引用を抽出

→引用関係のある9743ノード / 105175辺を取得

例) RFC9110 (HTTP Semantics)

9110→9111, 9110→1950, 9110→1951, …の有向辺を定義

19. References

19.1. Normative References

[CACHING] Fielding, R., Ed., Nottingham, M., Ed., and J. Reschke, Ed., "HTTP Caching", STD 98, RFC 9111, DOI 10.17487/RFC9111, June 2022, <<https://www.rfc-editor.org/info/rfc9111>>.

[RFC1950] Deutsch, P. and J-L. Gailly, "ZLIB Compressed Data Format Specification version 3.3", RFC 1950, DOI 10.17487/RFC1950, May 1996, <<https://www.rfc-editor.org/info/rfc1950>>.

[RFC1951] Deutsch, P., "DEFLATE Compressed Data Format Specification version 1.3", RFC 1951, DOI 10.17487/RFC1951, May 1996, <<https://www.rfc-editor.org/info/rfc1951>>.

<https://datatracker.ietf.org/doc/html/rfc9110#name-references>

方法 (PageRank)

PageRank :

- 元はWebページのランキングアルゴリズム
- 被引用数 + 引用元の重要度 を合わせて各RFCの重要度を判定
 - 多く引用されているRFCほどスコアが高い
 - スコアが高いRFCから引用されているほどスコアが高い

ツール :

- NetworkX の pagerank を $\alpha=0.85$ として使用

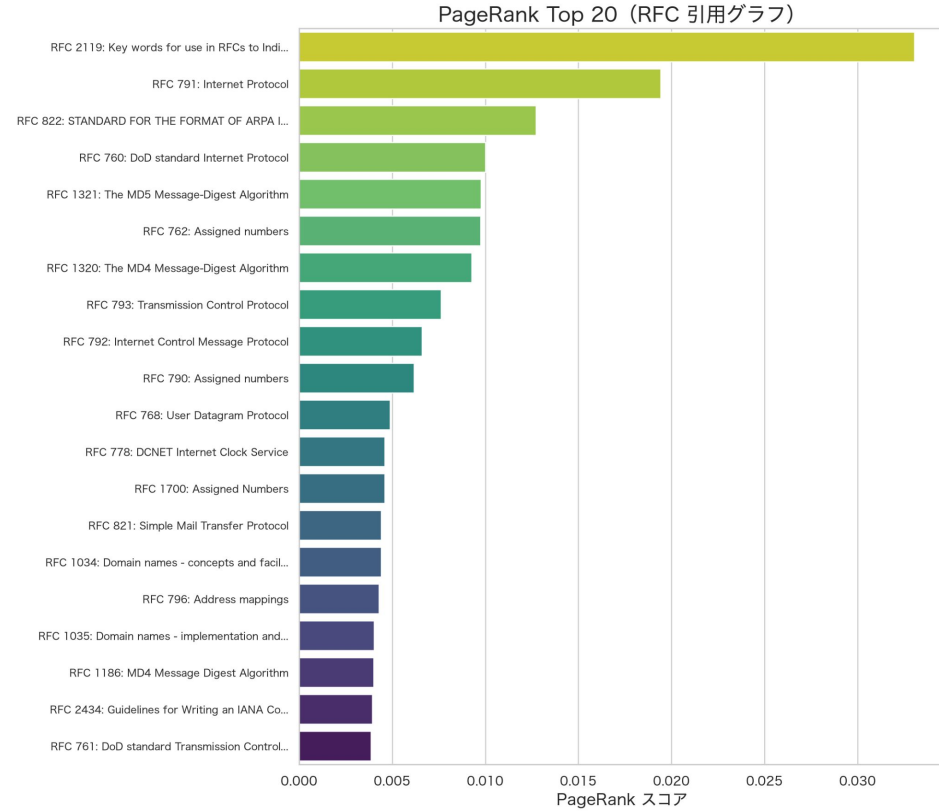
結果①

1位: RFC 2119
(RFCでのMUST, SHOULDの役割)

2位: RFC 791 (IPv4)

3位: RFC 822
(インターネットメールメッセージ形式)

古いRFCが強い！



結果②

2010年以降のRFCを
被引用数/経過年数 でランキング

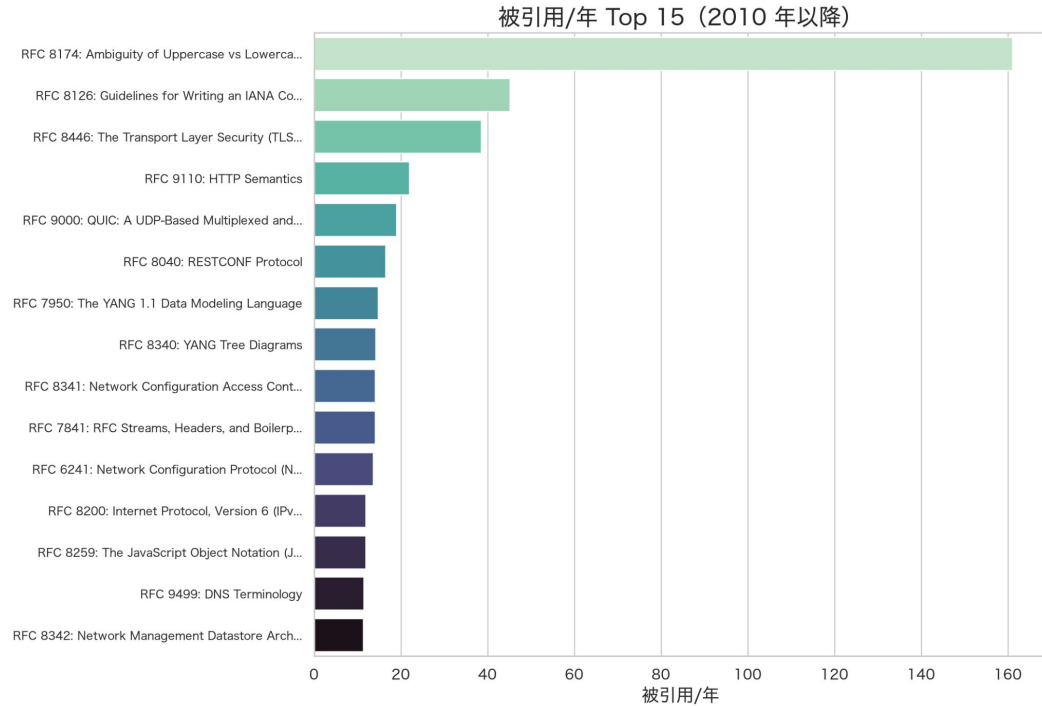
1位: RFC 8174
(RFC 2119をUpdates)

2位: RFC 8126
(IANA Considerationsの書き方)

3位: RFC 8446 (TLS 1.3)

4位: RFC 9110 (HTTP Semantics)

5位: RFC 9000 (QUIC)



結論（RFCの読み方の候補）

Step 1 - RFCの読み方

- RFC 2119 / RFC 8174 (BCP 14)
→MUST/SHOULD/MAYの意味

Step 2 - 引用の伸びが大きいプロトコル系（2010年以降 被引用/年 3~5位）

- RFC 8446 (TLS 1.3)
- RFC 9110 (HTTP Semantics)
- RFC 9000 (QUIC)

現状の手法の限界

- statusをみていない
(pagerank3位のRFC 822は、RFC 2822にobsolete→現行はRFC 5322となっている、等)
- 重みなしの有向グラフ
(Normative Reference, Informative Referenceの区別をしていない)
- 分野別グラフを作っていない
(SNMP等、部分グラフの方が「読む候補」として扱いやすい)

Enabling a Connected Future.



BBSakuraNetworks