

IETF124 MONTREAL

Automatic Mining of Mailing Lists for Internet Protocol Design Decisions

How does standardisation look like in an LLM-enabled future

Jie Bian
Network and Distributed System
University of Oslo

INFINITE POSSIBILITIES BETWEEN IETF AND LLM

- IETF has rich textual resources, including Internet-Drafts, RFCs, Email Archive and WG GitHub repositories, etc.
- Sources:
 - Datatracker (<https://datatracker.ietf.org>)
 - I-Ds and RFCs
 - Mailarchive (<https://mailarchive.ietf.org>)
 - Emails
 - GitHub
 - 63 WGs with 215 draft repositories found by Dec. 2024;
 - 159 new draft repositories (from some of the above WGs and new ones) by Sept. 2025

MOTIVATION: UNDERSTANDING DESIGN DECISIONS

- RFCs serve as the “cook books” for implementers.
 - They might or might not give some explanation for why things are the way they are.
 - For education or for better design, understanding the “*why*” is crucial!
 - The “*why*” information lie somewhere else—the Email Archive, Github, meeting minutes.
 - We aim to find the “*why*” from the email archive: decisions that led to RFC text.
 - RFC8374 is the only official document we found that documents design decisions (of RFC8205) and the corresponding rationale discussed in the emails.
-

HOW WE UTILIZE IETF DATA: WHAT WE HAVE DONE AND WHERE WE ARE NOW

- We manually annotated emails at first:
 - 2 RFCs from 2 WGs. Time-consuming - we can not expand the work to more RFCs or WGs
- We then moved to GitHub Repositories to gain more data:
 - Scale: 63 WGs with 215 draft repositories by Dec. 2024;
 - Automatic dataset labels via GitHub's encoded Pull Request (PR) – Issue relationship.
- **Now we attempt to leverage the Email Archive:**
 - We've collected emails from 63 WGs.
 - We asked LLM to “recover” design decision discussions in emails (generating synthetic data using LLM and prompts), and we create “RFCAlign”, a semi-synthetic dataset.

RFC8374 (RFC8205)

- **Decision:**

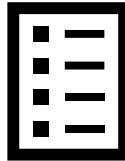
Withdrawals are not signed.

- **Discussion:**

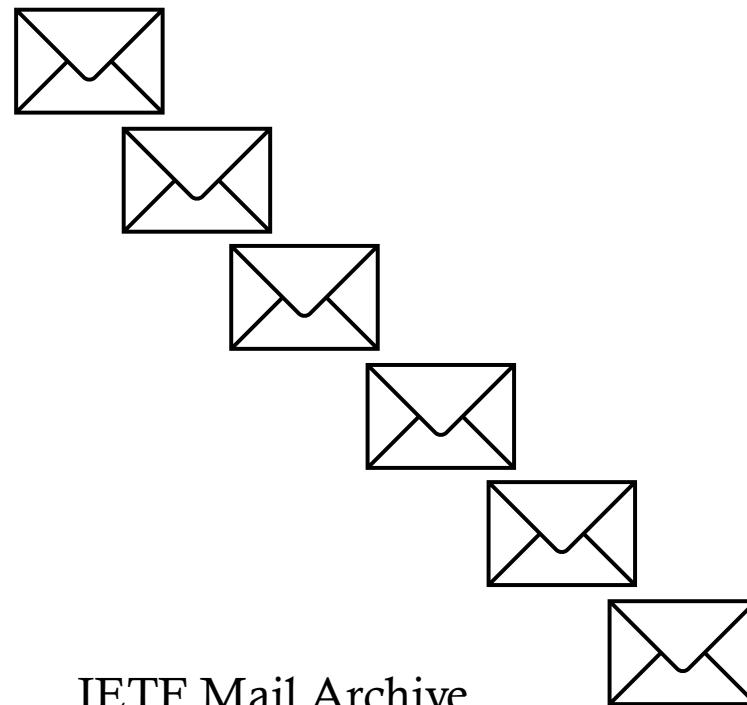
In the current BGP protocol, any AS can withdraw, at any time, any prefix it previously announced. *The rationale for not signing withdrawals is that BGPsec assumes the use of transport security between neighboring BGPsec routers.* Thus, no external entity can inject an update that withdraws a route or replay a previously transmitted update containing a withdrawal...

RELEVANT DISCUSSION LIES IN THE EMAIL ARCHIVE

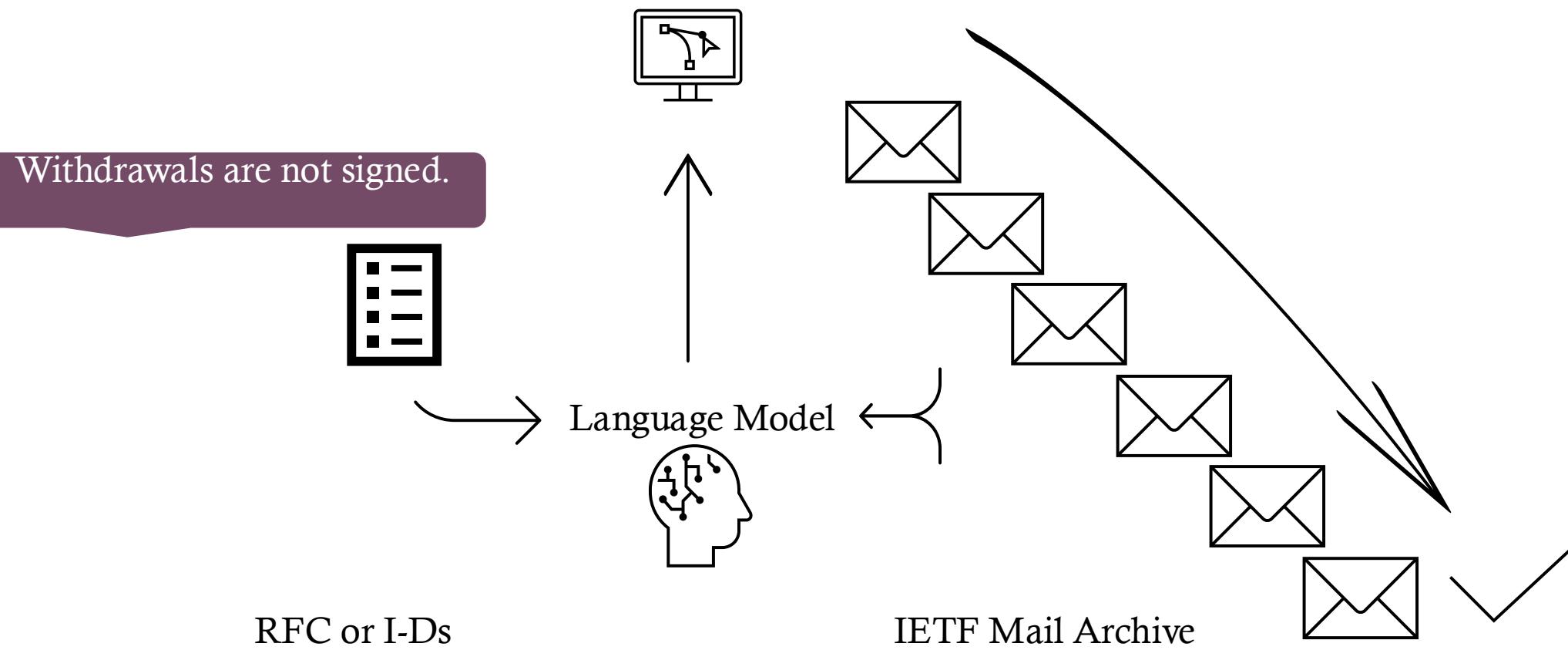
Withdrawals are not signed.



RFC or I-Ds



OUR METHODOLOGY: INFORMATION RETRIEVAL



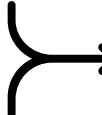
EXAMPLE FROM OUR GITHUB DATA: RETRIEVE RELEVANT DISCUSSION

Before:

If no ACK_FREQUENCY frames have been received, this value **defaults to 3**, which is the recommended packet threshold for loss detection in (Section 18.2 of QUIC-RECOVERY).

After:

If no ACK_FREQUENCY frames have been received, the endpoint immediately acknowledges any subsequent packets that are received out of order, as specified in Section 13.2 of QUIC-TRANSPORT, as such the **default value is 1**.



Discussion/Comment:

It seems optimal to default it to 3, because that's the default packet threshold, but the principle of least surprise argues for 1 (ie: no behavior change until an ACK_FREQUENCY frame is received), so I lean toward 1.

EXAMPLE FROM OUR GITHUB DATA: LLM AUTO EDIT

Before:

If no ACK_FREQUENCY frames have been received, this value **defaults to 3**, which is the recommended packet threshold for loss detection in (Section 18.2 of QUIC-RECOVERY).

After: Ground Truth

If no ACK_FREQUENCY frames have been received, the endpoint immediately acknowledges any subsequent packets that are received out of order, as specified in Section 13.2 of QUIC-TRANSPORT, as such the **default value is 1**.



Discussion/Comments:

It seems optimal to default it to 3, because that's the default packet threshold, but the principle of least surprise argues for 1 (ie: no behavior change until an ACK_FREQUENCY frame is received), so I lean toward 1.

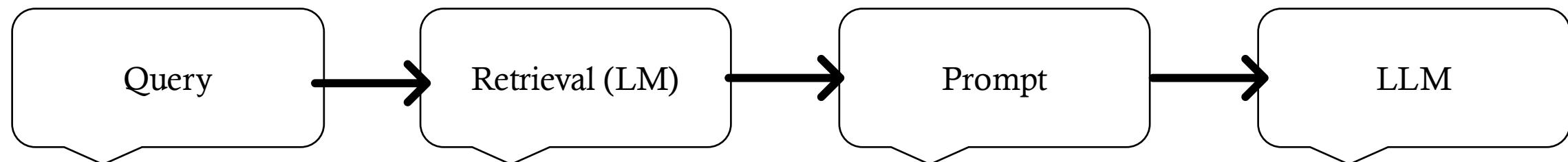
Edited by LLM:

If no ACK_FREQUENCY frames have been received, this value **defaults to 1**, which aligns with the principle of least surprise, maintaining the status quo until an ACK_FREQUENCY frame is received.

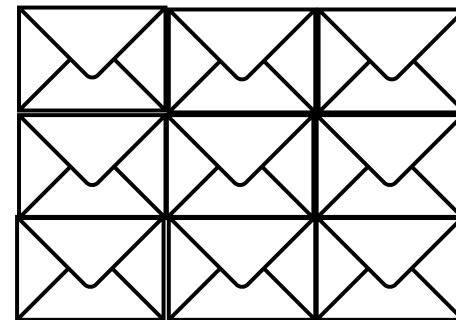
EXPAND FROM GITHUB TO EMAIL ARCHIVE (WORK UNDER SUBMISSION)

- RFCs defines the decisions but does not include the rationale.
- Emails include the rationale but may or may not mention the decision.
- We ask the LLM to restore the missing decisions discussed in the email threads, as synthetic data, and validate the effectiveness of our approach on the email dataset.

RETRIEVAL AUGMENTED GENERATION (FUTURE WORK)



Withdrawals are not signed.



Based on the retrieved discussion, explain why...



The rationale for not signing withdrawals is that BGPsec assumes the use of transport security between neighboring BGPsec routers.

THANK YOU

- Our fine-tuned models and curated datasets are available on Huggingface : <https://huggingface.co/jiebi>
- Publications
 - IEEE ICMLCN 2024: "Tell Me Why: Language Models Help Explain the Rationale Behind Internet Protocol Design"
 - We define a classification task where the LM (BERT) determine if a decision is relevant to a comment.
 - IEEE Access 2025: Automated Insights into GitHub Collaboration Dynamics
 - We leverage dense IR to search the relevant issue comments for certain design changes, or the other way around.
 - NLDB 2025: Instruction Tuning TextFlow Semi-Automatic RFCs Generation
 - Maximizing the potential of IETF data from a natural language processing (NLP) perspective
 - ACM ANRW 2025: Empowering IETF Collaboration with NLP Search Innovations and LLM-Enhanced RFC Writing
 - We utilize LLM to auto-edit RFC, in addition to searching