L-017

暗号解析プロセスの自動化を実現する生成 AI プロンプトに関する一考察

A Study on Generative AI Prompts for Automating the Cryptanalytic Process

杉尾 信行 †
北海道科学大学 †

1 はじめに

軽量暗号に対する代表的な攻撃手法として，差分攻撃 [1] や線形攻撃 [2] が広く知られている．これらの攻撃を行うにあたり，攻撃者は対象とする暗号の構造から攻撃に利用可能な「識別子」を事前に探索する必要がある．近年では，この識別子探索の効率化のため，混合整数線形問題（MILP）や充足可能問題（SAT）のソルバーを活用する手法が多数提案されている [3, 4, 5, 6].

MILP や SAT を用いた識別子探索は，軽量暗号の安全性評価における強力なツールとなっている．しかしながら，これらの手法を実際に適用するには，対象暗号の構造を数学的な制約式として記述するための専門的な知識に加え，ソルバーを扱うためのプログラミングスキルが求められる．この専門性が，研究者や技術者が軽量暗号の安全性評価に取り組む上での障壁となっている．

そこで本研究では，生成 AI 技術を活用することで，軽量暗号の安全性評価プロセスにおける上記障壁の低減，ひいてはその自動化の可能性を探ることを目指す．本稿では，MILP や SAT を用いた暗号解析プログラムの作成を支援する生成 AI プロンプトの設計に関する初期的な検討結果を報告する．

2 関連研究

2.1 共通鍵暗号の実装における生成 AI の応用

大規模言語モデル（LLM）を共通鍵暗号分野に応用した先行研究として，暗号アルゴリズムの実装に関する試みが挙げられる．Kwon ら [7] や Cintas-Canto ら [8] は，ChatGPT を用いて共通鍵暗号 AES，CHAM，ASCON などのプログラムコードを生成し，その実現可能性を示した．

2.2 プログラムの実装における生成 AI の応用

著者らは，自身の研究において，暗号解析プログラムの実装に生成 AI の ChatGPT を活用する試みを実施している [9]．2024 年 1 月時点で実施したこの取り組みを通じて，以下の課題が明らかになった．

- 生成されたコードが，対象とする暗号アルゴリズムに存在しない関数や構造を参照するなど，事実に基づかない内容を含むことがある（ハルシネーション）．
- S-box などの特定の暗号コンポーネントの解析に特化した外部アプリケーション（Logic Friday など）で得られた結果を，生成 AI に引き継いでプログラム生成を行うといったツール連携が難しい．

3 生成 AI プロンプト設計の検討

前述の課題を克服し，MILP や SAT を用いた暗号解析プログラムの生成精度を高めるためには，生成 AI への指示である「プロンプト」の設計を改良することが有効と考えられる．本章では，そのための具体的な検討事項を述べる．

3.1 ハルシネーション対策

生成 AI によるハルシネーションを抑制し，正確なコードを生成させるために，プロンプトには以下の情報を明確かつ詳細に含めることが有効であると思われる．

- 解析対象となる暗号アルゴリズムの正式名称，バージョン，構成（ラウンド数，ブロックサイズ，鍵サイズなど）を正確に指定する．
- 暗号を構成する主要な要素（S-box，線形変換層など）について，その定義を厳密に記述する．例えば，S-box の場合は入出力テーブル，線形変換の場合はビット単位の操作や数学的な行列表現などを具体的に提供する．

3.2 S-box の制約式生成

暗号解析に不可欠な S-box の特性（差分特性や線形特性など）を表す制約式を生成させる際には，いくつかの方法が考えられる．ここでは，生成 AI に S-box の解析自体を行わせる方法（方法 1）と，外部ツールによる解析結果を生成 AI に与える方法（方法 2）について，プロンプト設計の観点から検討する．

- **方法 1（生成 AI による解析と制約式導出）：** S-box の特性を SAT または MILP の制約式として表現させる場合，制約式における変数の定義方法や，論理式・線形不等式の記述ルールなどを具体的に指示する必要がある．例えば，「この S-box の入出力差分（または入出力マスク）間の関係を表す論理

† Nobuyuki Sugio, Hokkaido University of Science

式（または線形不等式）を生成し，それを標準的な SAT（CNF）または MILP 形式で出力してください」といった具体的な指示に加えて，使用する変数の命名規則や意味，および論理積や論理和，不等号などの数学的・論理的記号の表現方法を詳細に定義することが重要となる．

- **方法 2（外部ツールによる解析結果に基づく制約式導出）：** S-box の解析を S-box analyzer[1]のような外部ツールで行い，その解析結果（例: 差分伝播確率が非ゼロとなる入出力差分のペアリスト）を生成 AI に与えて制約式のみを生成させる場合も考えられる．この場合も，SAT または MILP における変数定義や論理式・線形不等式の記述ルールをプロンプト内で明確に示すことが不可欠である．

4 一般的考慮事項

上記の暗号解析固有の課題対策に加え，生成 AI による高品質なコード生成を促すための一般的なプロンプト設計のベストプラクティスも，暗号解析プログラムの生成に適用できると考えられる．

- **明確性・網羅性:** 生成してほしいコードが「何を（解析対象の暗号，攻撃手法）」「どのように（MILP ソルバーまたは SAT ソルバーを利用）」「どのプログラミング言語で（Python など）」「どのような形式で（コードの全体構造，関数や変数の命名規約，コメントの付加など）」出力されるべきかを，曖昧さなく明確に伝える．また，使用を想定している特定のライブラリやフレームワーク（例: MILP ソルバーの `Gurobi Python API`，SAT ソルバーの `PySAT` など）があれば，それを指定する．
- **段階的な指示:** 一度に複雑なプログラム全体を生成させようとするのではなく，暗号の 1 ラウンド分のモデル記述，ラウンド間の接続方法，目的関数の設定，特定の探索条件の追加など，プログラムを構成する要素ごとに分けてプロンプトを作成し，段階的にコードを生成させる．
- **例示:** 可能な限り，生成してほしいコードの構造やスタイルを示す短いコード断片や，期待される出力形式の具体的な例をプロンプトに含める．これにより，生成 AI の出力がより意図に沿ったものになる可能性が高まる．

5 まとめ

本稿では，生成 AI を暗号解析プログラムの作成に活用する際に顕在化した課題に対し，その克服に向けた生成 AI プロンプト設計に関する初期的な検討を行った．具体的には，生成 AI のハルシネーション抑制策と，外部ツール連携が課題となる S-box の制約式生成に関するプロンプト設計の方向性を示した．また，効果的なコード生成のための一般的なプロンプト設計の考慮事項についても触れた．ハルシネーションや外部ツールとの連携といった実用上の課題を実際に克服し，生成 AI を用いた暗号解析プロセスの自動化を実現するためには，本稿で検討したプロンプト設計指針に基づき，具体的なプロンプトを作成し，生成されるコードの精度や信頼性を定量的に評価していくことが不可欠である．これが今後の研究における重要なステップとなる．

[1] `https://github.com/hadipourh/sboxanalyzer`

謝辞

本研究の一部は，JSPS 科研費 JP25K15119 の助成を受けたものである．

参考文献

[1] E. Biham, A. Shamir. “Differential Cryptanalysis of the Data Encryption Standard”, Springer-Verlag, New York, pp.79-88, 1993.

[2] M. Matsui. “Linear Cryptanalysis Method for DES Cipher”, Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques (EUROCRYPT ’93), LNCS 765, pp.386-397, 1993.

[3] N. Mouha, Q. Wang, D. Gu and B. Preneel, “Differential and Linear Cryptanalysis Using Mixed-Integer Linear Programming”, in Proceeding of the International Conference on Information Security and Cryptology, Inscrypt 2011, vol.7537 of LNCS, pp.57-76, Springer-Verlag, 2011.

[4] Y. Sasaki, Y. Todo, New “Impossible Differential Search Tool from Design and Cryptanalysis Aspects”, Proceeding of the 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, EUROCRYPT 2017, vol.10212 of LNCS, pp.185–215, Springer-Verlag, 2017.

[5] L. Sun, W. Wang, M. Wang, “More Accurate Differential Properties of LED64 and Midori64”, IACR Transactions on Symmetric Cryptology, vol.2018, Issue.3, no.3, pp.93-123, 2018.

[6] L. Sun, W. Wang, M. Wang, “Accelerating the Search of Differential and Linear Characteristics with the SAT Method”, IACR Transactions on Symmetric Cryptology, vol.2021, Issue.1, no.1, pp.269-315, 2021.

[7] H. Kwon, M. Sim, G. Song, M. Lee, H. Seo, “Novel Approach to Cryptography Implementation using ChatGPT”, IACR Cryptology ePrint Archive: Report 2023/606, 2023.

[8] A. Cintas-Canto, J. Kaur, M. Mozaffari-Kermani, R. Azarderakhsh, “ChatGPT vs. Lightweight Security: First Work Implementing the NIST Cryptographic Standard ASCON”, arXiv:2306.08178, 2023.

[9] N. Sugio.: Implementation of Cryptanalytic Program for ASCON Using ChatGPT, Proc. 2024 Twelfth International Symposium on Computing and Networking Workshops (CANDARW), pp. 307-313, 2024.