Webacy Risk Tags Correlation and Frequency

Project by Ben (Minh) Pham

Overview

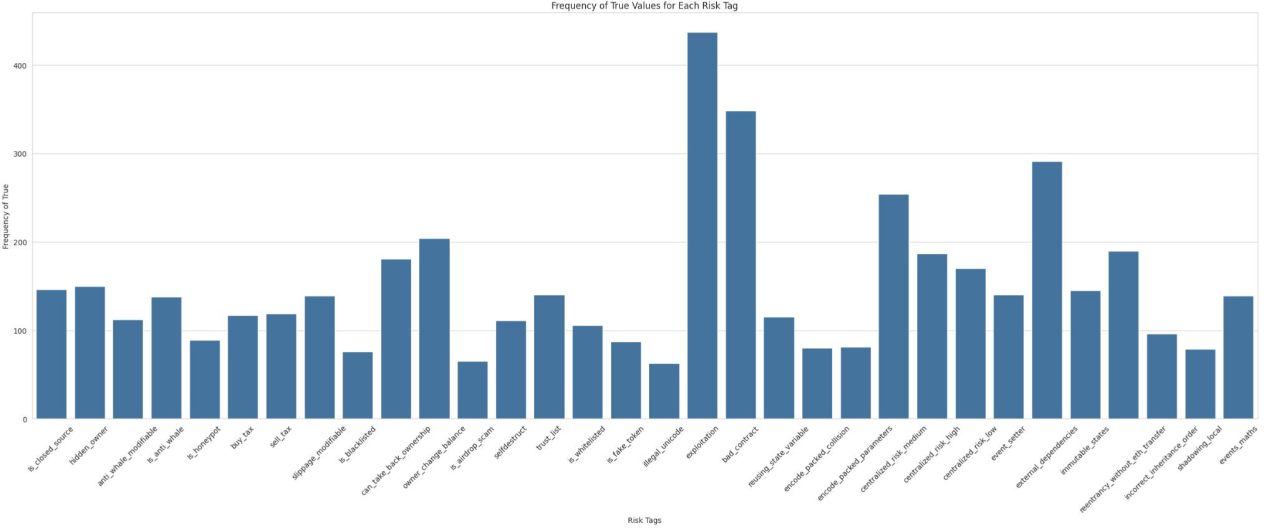
- Introduction
- Frequency Analysis
- Heatmap of Correlations
- Strong Positive Correlations
- Surprising Findings | Negative Correlations
- Conclusion

Introduction

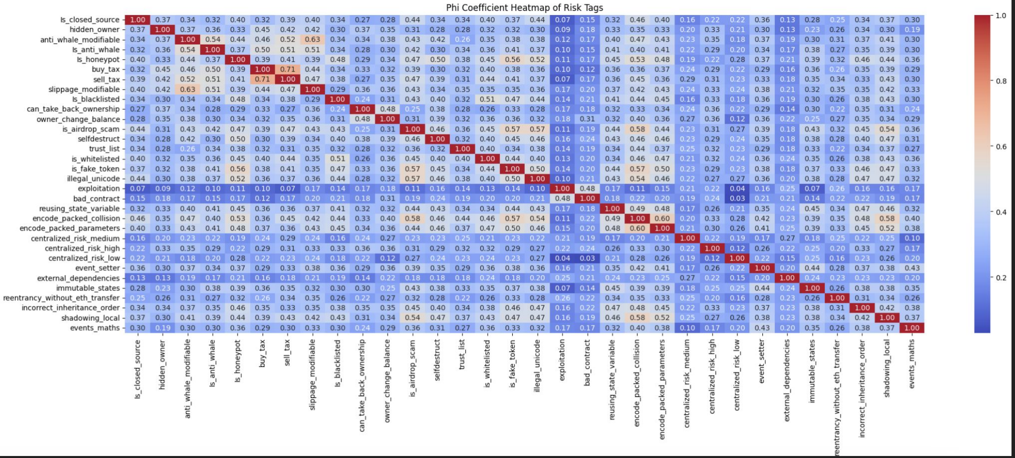
- Dataset: Risk tags associated with smart contract risks Vulnerabilities.
- Objective: Analyze the frequency as well as the correlation among risk tags.
- Relevance: This makes secure smart contracts indispensable in decentralized finance. It must be mentioned that recognizing risk patterns assists in enhancing the contract’s safety.

Frequency Analysis

- Top Risk Tags: exploitation appears most frequently, while bad_contract is second.
- Key Insights: Exploitation risks are highly prevalent in smart contracts.
- Implications: Companies need robust security audits to mitigate this risk.



Heatmap of Correlations

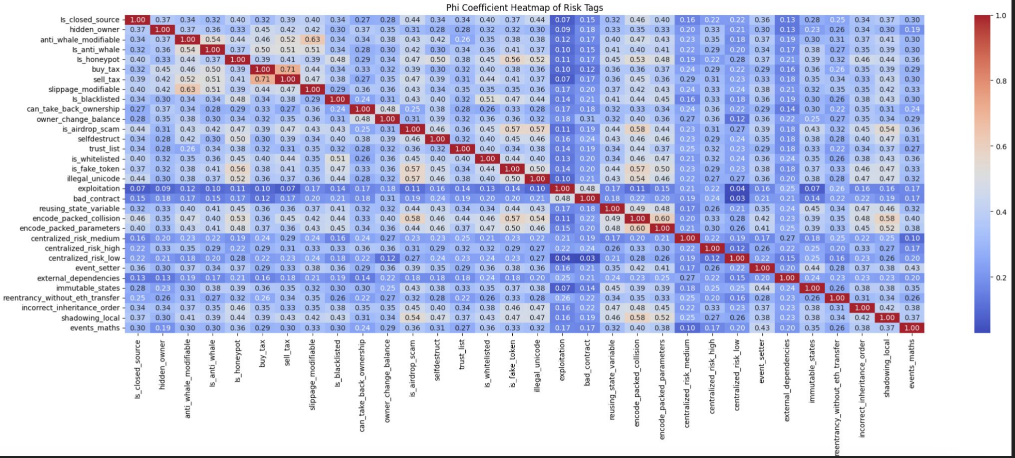


Phi Coefficient Heatmap of Risk Tags

	is_closed_source	hidden_owner	anti_whale_modifiable	is_anti_whale	is_honeypot	buy_tax	sell_tax	slippage_modifiable	is_blacklisted	can_take_back_ownership	owner_change_balance	is_airdrop_scam	selfdestruct	trust_list	is_whitelisted	is_fake_token	illegal_unicode	exploitation	bad_contract	reusing_state_variable	encode_packed_collision	encode_packed_parameters	centralized_risk_medium	centralized_risk_high	centralized_risk_low	event_setter	external_dependencies	immutable_states	reentrancy_without_eth_transfer	incorrect_inheritance_order	shadowing_local	events_maths
is_closed_source	1.00	0.37	0.34	0.32	0.40	0.32	0.39	0.40	0.34	0.27	0.28	0.44	0.34	0.34	0.40	0.37	0.44	0.07	0.15	0.32	0.46	0.40	0.16	0.22	0.22	0.36	0.13	0.28	0.25	0.34	0.37	0.30
hidden_owner	0.37	1.00	0.37	0.36	0.33	0.45	0.42	0.42	0.30	0.37	0.35	0.31	0.28	0.28	0.32	0.32	0.30	0.09	0.18	0.33	0.35	0.33	0.20	0.33	0.21	0.30	0.13	0.23	0.26	0.34	0.30	0.19
anti_whale_modifiable	0.34	0.37	1.00	0.54	0.44	0.46	0.52	0.63	0.34	0.34	0.38	0.43	0.42	0.26	0.35	0.38	0.38	0.12	0.17	0.40	0.47	0.43	0.23	0.35	0.18	0.37	0.19	0.30	0.31	0.37	0.41	0.30
is_anti_whale	0.32	0.36	0.54	1.00	0.37	0.50	0.51	0.51	0.34	0.28	0.30	0.42	0.30	0.34	0.36	0.41	0.37	0.10	0.15	0.41	0.40	0.41	0.22	0.29	0.20	0.34	0.17	0.38	0.27	0.35	0.39	0.30
is_honeypot	0.40	0.33	0.44	0.37	1.00	0.39	0.41	0.39	0.48	0.29	0.34	0.47	0.50	0.38	0.45	0.56	0.52	0.11	0.17	0.45	0.53	0.48	0.19	0.22	0.28	0.37	0.21	0.39	0.32	0.46	0.44	0.36
buy_tax	0.32	0.45	0.46	0.50	0.39	1.00	0.71	0.44	0.34	0.33	0.32	0.39	0.30	0.32	0.40	0.38	0.36	0.10	0.12	0.36	0.36	0.37	0.24	0.29	0.22	0.29	0.16	0.36	0.26	0.35	0.39	0.29
sell_tax	0.39	0.42	0.52	0.51	0.41	0.71	1.00	0.47	0.38	0.27	0.35	0.47	0.39	0.31	0.44	0.41	0.37	0.07	0.17	0.36	0.45	0.36	0.29	0.31	0.23	0.33	0.18	0.35	0.34	0.33	0.43	0.30
slippage_modifiable	0.40	0.42	0.63	0.51	0.39	0.44	0.47	1.00	0.29	0.36	0.36	0.43	0.34	0.35	0.35	0.35	0.36	0.17	0.20	0.37	0.42	0.43	0.24	0.33	0.24	0.38	0.21	0.32	0.35	0.35	0.42	0.33
is_blacklisted	0.34	0.30	0.34	0.34	0.48	0.34	0.38	0.29	1.00	0.24	0.31	0.43	0.40	0.32	0.51	0.47	0.44	0.14	0.21	0.41	0.44	0.45	0.16	0.33	0.18	0.36	0.19	0.30	0.26	0.38	0.43	0.30
can_take_back_ownership	0.27	0.37	0.34	0.28	0.29	0.33	0.27	0.36	0.24	1.00	0.48	0.25	0.38	0.28	0.26	0.33	0.28	0.17	0.18	0.32	0.33	0.34	0.24	0.36	0.22	0.29	0.14	0.30	0.22	0.35	0.31	0.24
owner_change_balance	0.28	0.35	0.38	0.30	0.34	0.32	0.35	0.36	0.31	0.48	1.00	0.31	0.39	0.32	0.36	0.36	0.32	0.18	0.31	0.32	0.40	0.36	0.27	0.36	0.12	0.36	0.22	0.25	0.27	0.35	0.34	0.29
is_airdrop_scam	0.44	0.31	0.43	0.42	0.47	0.39	0.47	0.43	0.43	0.25	0.31	1.00	0.46	0.36	0.45	0.57	0.57	0.11	0.19	0.44	0.58	0.44	0.23	0.31	0.27	0.39	0.18	0.43	0.32	0.45	0.54	0.36
selfdestruct	0.34	0.28	0.42	0.30	0.50	0.30	0.39	0.34	0.40	0.38	0.39	0.46	1.00	0.32	0.40	0.45	0.46	0.16	0.24	0.43	0.46	0.46	0.23	0.29	0.24	0.35	0.18	0.38	0.28	0.40	0.47	0.31
trust_list	0.34	0.28	0.26	0.34	0.38	0.32	0.31	0.35	0.32	0.28	0.32	0.36	0.32	1.00	0.40	0.34	0.38	0.14	0.19	0.34	0.44	0.37	0.25	0.32	0.23	0.29	0.18	0.33	0.22	0.34	0.37	0.27
is_whitelisted	0.40	0.32	0.35	0.36	0.45	0.40	0.44	0.35	0.51	0.26	0.36	0.45	0.40	0.40	1.00	0.44	0.40	0.13	0.20	0.34	0.46	0.47	0.21	0.32	0.24	0.36	0.24	0.35	0.26	0.38	0.43	0.36
is_fake_token	0.37	0.32	0.38	0.41	0.56	0.38	0.41	0.35	0.47	0.33	0.36	0.57	0.45	0.34	0.44	1.00	0.50	0.14	0.20	0.44	0.57	0.50	0.23	0.29	0.23	0.38	0.18	0.37	0.33	0.46	0.47	0.33
illegal_unicode	0.44	0.30	0.38	0.37	0.52	0.36	0.37	0.36	0.44	0.28	0.32	0.57	0.46	0.38	0.40	0.50	1.00	0.10	0.21	0.43	0.54	0.46	0.22	0.27	0.27	0.36	0.20	0.38	0.28	0.47	0.47	0.32
exploitation	0.07	0.09	0.12	0.10	0.11	0.10	0.07	0.17	0.14	0.17	0.18	0.11	0.16	0.14	0.13	0.14	0.10	1.00	0.48	0.17	0.11	0.15	0.21	0.22	0.04	0.16	0.25	0.07	0.26	0.16	0.16	0.17
bad_contract	0.15	0.18	0.17	0.15	0.17	0.12	0.17	0.20	0.21	0.18	0.31	0.19	0.24	0.19	0.20	0.20	0.21	0.48	1.00	0.18	0.22	0.20	0.19	0.24	0.03	0.21	0.21	0.14	0.22	0.22	0.19	0.17
reusing_state_variable	0.32	0.33	0.40	0.41	0.45	0.36	0.36	0.37	0.41	0.32	0.32	0.44	0.43	0.34	0.34	0.44	0.43	0.17	0.18	1.00	0.49	0.48	0.17	0.26	0.21	0.35	0.24	0.45	0.34	0.47	0.46	0.32
encode_packed_collision	0.46	0.35	0.47	0.40	0.53	0.36	0.45	0.42	0.44	0.33	0.40	0.58	0.46	0.44	0.46	0.57	0.54	0.11	0.22	0.49	1.00	0.60	0.20	0.33	0.28	0.42	0.23	0.39	0.35	0.48	0.58	0.40
encode_packed_parameters	0.40	0.33	0.43	0.41	0.48	0.37	0.36	0.43	0.45	0.34	0.36	0.44	0.46	0.37	0.47	0.50	0.46	0.15	0.20	0.48	0.60	1.00	0.21	0.30	0.26	0.41	0.25	0.39	0.33	0.45	0.52	0.38
centralized_risk_medium	0.16	0.20	0.23	0.22	0.19	0.24	0.29	0.24	0.16	0.24	0.27	0.23	0.23	0.25	0.21	0.23	0.22	0.21	0.19	0.17	0.20	0.21	1.00	0.22	0.19	0.17	0.27	0.18	0.25	0.22	0.25	0.10
centralized_risk_high	0.22	0.33	0.35	0.29	0.22	0.29	0.31	0.33	0.33	0.36	0.36	0.31	0.29	0.32	0.32	0.29	0.27	0.22	0.24	0.26	0.33	0.30	0.22	1.00	0.12	0.26	0.22	0.25	0.20	0.33	0.27	0.17
centralized_risk_low	0.22	0.21	0.18	0.20	0.28	0.22	0.23	0.24	0.18	0.22	0.12	0.27	0.24	0.23	0.24	0.23	0.27	0.04	0.03	0.21	0.28	0.26	0.19	0.12	1.00	0.22	0.15	0.25	0.16	0.23	0.26	0.20
event_setter	0.36	0.30	0.37	0.34	0.37	0.29	0.33	0.38	0.36	0.29	0.36	0.39	0.35	0.29	0.36	0.38	0.36	0.16	0.21	0.35	0.42	0.41	0.17	0.26	0.22	1.00	0.20	0.44	0.28	0.37	0.38	0.43
external_dependencies	0.13	0.13	0.19	0.17	0.21	0.16	0.18	0.21	0.19	0.14	0.22	0.18	0.18	0.18	0.24	0.18	0.20	0.25	0.21	0.24	0.23	0.25	0.27	0.22	0.15	0.20	1.00	0.24	0.23	0.23	0.23	0.20
immutable_states	0.28	0.23	0.30	0.38	0.39	0.36	0.35	0.32	0.30	0.30	0.25	0.43	0.38	0.33	0.35	0.37	0.38	0.07	0.14	0.45	0.39	0.39	0.18	0.25	0.25	0.44	0.24	1.00	0.26	0.38	0.38	0.35
reentrancy_without_eth_transfer	0.25	0.26	0.31	0.27	0.32	0.26	0.34	0.35	0.26	0.22	0.27	0.32	0.28	0.22	0.26	0.33	0.28	0.26	0.22	0.34	0.35	0.33	0.25	0.20	0.16	0.28	0.23	0.26	1.00	0.31	0.34	0.26
incorrect_inheritance_order	0.34	0.34	0.37	0.35	0.46	0.35	0.33	0.35	0.38	0.35	0.35	0.45	0.40	0.34	0.38	0.46	0.47	0.16	0.22	0.47	0.48	0.45	0.22	0.33	0.23	0.37	0.23	0.38	0.31	1.00	0.42	0.38
shadowing_local	0.37	0.30	0.41	0.39	0.44	0.39	0.43	0.42	0.43	0.31	0.34	0.54	0.47	0.37	0.43	0.47	0.47	0.16	0.19	0.46	0.58	0.52	0.25	0.27	0.26	0.38	0.23	0.38	0.34	0.42	1.00	0.37
events_maths	0.30	0.19	0.30	0.30	0.36	0.29	0.30	0.33	0.30	0.24	0.29	0.36	0.31	0.27	0.36	0.33	0.32	0.17	0.17	0.32	0.40	0.38	0.10	0.17	0.20	0.43	0.20	0.35	0.26	0.38	0.37	1.00

- Heatmap Colors: The colors vary from one extreme to the other (often from blue to red in a 'coolwarm' palette), showing the degree and direction of the link.
- In terms of the Phi coefficient:
  - Darker reds indicate stronger positive correlations, implying that if one risk tag is true, the other is more likely to be true.
  - Darker blues show stronger negative correlations, implying that if one risk tag is accurate, the other is likely to be incorrect.
  - Neutral hues (near white or the middle of the palette) indicate little to no link between the risk markers.
- The higher the correlation the redder it is and the lower the bluer it is.

Strong Positive Correlations



- Phi-Coefficient: Strong correlation (0.71) between buy_tax and sell_tax.
- Phi-Coefficient: Strong correlation (0.60) between encoded_packed_collisions and encoded_packed_parameters.
- Explanation: These tags can be seen frequently in combined with tokenomics models. Then for the encoded both risks are related to how data is packed and passed within smart contracts.
- Actionable Insight: Both therefore needs to be assessed by auditors when assessing tax-based token contracts. Then for the encoded they need to be fixed by having better coding practices.

Surprising Findings | Negative Correlations

- hidden_owner and is_airdrop_scam only has a correlation of 0.28. This could be because they could be phishing or using social engineering technique for the scam.
- Is_whitelisted and trustlist only has a correlation of 0.40. Trust list might be only used in niche scenarios allowing for permissions in certain scenarios.
- centralized_low_risk and exploitation has a correlation of only 0.04.
- Bad_contract and centralized_low_risk only has a correlation of 0.03.
- Is_closed_source and exploitation only has a correlation of 0.07

Summary:

- To this end, this presentation seeks to analyze the risk tags of smart contract vulnerabilities and more particularly their distribution and co-occurrence.
- Key findings include:
    - Top Risk Tags: Exploitation is the most often related risk which emphasizes the importance of proper security checks for smart contracts.
    - Correlation Analysis: This is shown by heatmap where the positive correlations like, between buy_tax and sell_tax as well as between encoded_packed_collisions and encoded_packed_parameter. Such correlations strengthen the call for thorough auditing in such areas as tokenomics and encoding of data.
    - Surprising Findings: Some presumed positive associations, as between hidden_owner and is_airdrop_scam, or between centralized_low_risk and exploitation, were less than expected. These results indicate the necessity of the subsequent research studies concerning these relations.
    - Thus, the created analysis emphasizes that the threats connected to smart contracts are diverse and require careful and specific approaches to be eliminated.

Thank You
By Ben Pham

webacy