

PREFIX VULNERABILITY METRICS — THEORY & SECURITY RATIONALE

Why These 7 Metrics Are Sufficient for Predicting Prefix-Level Routing Risk

1. Introduction

Prefix vulnerability in BGP refers to the susceptibility of an IP prefix to being hijacked, misdirected, substituted, or leaked due to weaknesses in origin validation, visibility, routing hygiene, or operator configuration.

The goal of this document is to justify — **rigorously and transparently** — why the seven selected metrics provide complete and sufficient coverage for assessing prefix-level security risk.

These metrics are:

1. **has_roa**
2. **vrp_count**
3. **roa_coverage_scope**
4. **roa_maxLength**
5. **irr_objects_count**
6. **prefix_length**
7. **vantage_points_seen**

Each metric corresponds to a distinct and necessary dimension of BGP security.

Together, they form a **360° vulnerability framework**.

2. What Is Prefix Vulnerability (Conceptually)?

A prefix is vulnerable when:

- It can be **originated from the wrong ASN**
- It can be **subprefix-hijacked** (/24 vs /23 etc.)
- It can be **propagated incorrectly** because filters are weak
- It is not **cryptographically verifiable**
- It has **incomplete or incorrect IRR configuration**
- It has **anomalous visibility** across RIS/route collectors

Prefix vulnerability therefore has **four core dimensions**:

1. **Origin authentication**
2. **Prefix specificity & propagation risk**
3. **RPKI/IRR integrity**
4. **Visibility & detectability**

These four dimensions (Origin Authentication, Specificity Risk, Routing Hygiene, Visibility) represent the entire academic and operational understanding of prefix vulnerability, as defined by IETF, CAIDA, RIPE, APNIC, MANRS, and major routing-security research institutions.

Note:

Vulnerability does not imply guaranteed exploitability — it represents susceptibility. Successful attacks require both a vulnerable prefix *and* a sufficiently vulnerable ASN. The platform's combined risk model quantifies exactly this interaction, identifying the most dangerous ASN+prefix combinations on the global Internet.

3. Rationale for Each Metric

3.1 ROA Presence — **has_roa**

What it measures:

Whether a cryptographic ROA exists for the prefix.

Why it matters:

Without a ROA, any ASN on Earth can originate the prefix.

Of course, the actual propagation success will depend on the ASN announcing it.

This is the biggest factor in prefix hijack vulnerability.

Security relevance:

- Prevents origin hijacks
- Provides cryptographic validation
- Required for ROV to work
- Without a ROA, the prefix is structurally unprotected and susceptible to origin hijacks.

Conclusion:

ROA presence is a foundational metric. Everything else builds on top of it.

3.2 VRP Count — **vrp_count**

What it measures:

How many valid ROA authorizations exist for a prefix.

Why it matters:

A prefix with 0–1 VRPs is fragile.

A prefix with many VRPs is strongly authorized.

Security relevance:

- Multi-VRP environments are more stable
- Reduces misconfiguration risk
- Indicates better operator hygiene
- Prefixes with few VRPs are often abandoned or poorly maintained

Conclusion:

VRP count gives depth and reliability to ROA authorization.

3.3 ROA Coverage Scope — **roa_coverage_scope**

What it measures:

How much of the prefix's announced address space is cryptographically authorized by ROAs (full, partial, or none).

Why it matters:

Many operators create ROAs *only for part* of their space.
Uncovered fragments are hijackable.

Security relevance:

- Identifies partial protection
- Measures misalignment between routing reality and ROAs
- Detects “gaps” attackers can exploit

Conclusion:

Full cryptographic origin authorization requires full coverage.
Partial coverage = security illusion.

3.4 ROA maxLength — **roa_maxLength**

What it measures:

The maximum allowed prefix length in the ROA.

Why it matters:

Misconfigured maxLength values are one of the most commonly exploited weaknesses in RPKI deployment.

If maxLength is too permissive, an attacker can announce valid subprefixes that ROV will accept, enabling high-impact subprefix hijacks.

This is a real-world failure frequently observed in BGP incidents.

Example:

A ROA for **10.0.0.0/16** with **maxLength=24** allows an attacker to hijack **10.0.0.0/24** even though the ROA “looks valid”.

Security relevance:

- Direct correlation to subprefix hijackability
- Detects permissive ROA policies
- Reveals structural misconfiguration

Conclusion:

No prefix vulnerability assessment is complete without maxLength.

3.5 IRR Object Count — **irr_objects_count**

What it measures:

How many authoritative IRR route objects exist for the prefix.

Why it matters:

IRR is imperfect but is still used by thousands of operators globally for filtering.

Prefixes with weak or inconsistent IRR records frequently belong to networks that:

- have outdated or incomplete IRR maintenance
- have administrative inconsistencies (missing, expired, or incorrect IRR route objects)
- lack automated IRR/RPKI synchronization workflows
- operate without strict routing policy enforcement

These conditions simply reflect the operational reality that many networks have limited automation or legacy processes.

Security relevance:

- Poor IRR ⇒ weak filtering
- Weak filtering ⇒ easier hijacks

- IRR inconsistencies \Rightarrow misconfigurations

Conclusion:

IRR is not perfect, but remains an essential filtering layer.

3.6 Prefix Length — **prefix_length**

What it measures:

CIDR size (/16, /17, /18.../24).

Why it matters:

Propagation and hijack behavior depend strongly on length:

- /24 is the minimum globally routable prefix in IPv4
- /25+ rarely propagate
- Large aggregates (/8–/15) are less likely to be hijacked successfully due to high visibility and strong monitoring, but hijacks are still technically possible.
- /22+/23+/24 are the most targeted

Security relevance:

- Determines propagation likelihood
- Determines how “attractive” the prefix is to attackers
- Reveals structural hijack potential

Conclusion:

Prefix length is fundamental to understanding vulnerability dynamics.

3.7 Vantage-Point Visibility — **vantage_points_seen**

What it measures:

How many RIS/route collector vantage points see the prefix.

Why it matters:

Low-visibility prefixes behave differently:

- are less monitored
- are easier to hijack without detection
- often originate from small / poorly-maintained networks

Security relevance:

- Indicates detectability
- Shows structural routing weakness
- Prefixes seen only in few vantage points are prime hijack targets

Conclusion:

Low visibility makes hijacks significantly harder to detect and correlates strongly with weak routing hygiene. Prefix visibility is therefore a critical vulnerability indicator.

4. Why These 7 Metrics Are Fully Sufficient

These metrics cover all **four pillars** of prefix security:

Security Dimension	Covered By
Origin Authentication	has_roa, vrp_count, roa_scope
Subprefix / Specificity Risk	roa_maxLength, prefix_length
Policy / Routing Hygiene	irr_objects_count
Visibility / Propagation	vantage_points_seen

Together, they form a comprehensive framework.

Nothing essential is missing.

There is **no fifth dimension** of prefix vulnerability that these metrics fail to capture.

5. Integration with ML

Each metric becomes an ML feature contributing to the final vulnerability score.

The ML engine learns patterns such as:

- prefixes without ROAs are high-risk
- prefixes with permissive maxLength are structurally vulnerable
- prefixes with low visibility often belong to unstable networks
- prefixes with poor IRR are linked to permissive operators
- combinations amplify risk nonlinearly

The ML does **not** rely on one metric alone — it learns interaction between features.

6. Conclusion

These seven metrics:

- are theoretically grounded
- are operationally meaningful
- represent all known dimensions of prefix vulnerability
- provide sufficient and complete coverage
- create a strong foundation for ML classification

- make the platform enterprise-grade and research-quality

This document proves that the prefix metrics are complete, sufficient, and rigorously justified .