# CWE Detail – CWE-185

## Description

The product specifies a regular expression in a way that causes data to be improperly matched or compared.

## Extended Description

When the regular expression is used in protection mechanisms such as filtering or validation, this may allow an attacker to bypass the intended restrictions on the incoming data.

## Threat-Mapped Scoring

Score: 0.0

Priority: Unclassified

## Observed Examples (CVEs)

**•** CVE-2002-2109: Regexp isn't "anchored" to the beginning or end, which allows spoofed values that have trusted values as substrings.

**•** CVE-2005-1949: Regexp for IP address isn't anchored at the end, allowing appending of shell metacharacters.

**•** CVE-2001-1072: Bypass access restrictions via multiple leading slash, which causes a regular expression to fail.

**•** CVE-2000-0115: Local user DoS via invalid regular expressions.

**•** CVE-2002-1527: chain: Malformed input generates a regular expression error that leads to information exposure.

**•** CVE-2005-1061: Certain strings are later used in a regexp, leading to a resultant crash.

**•** CVE-2005-2169: MFV. Regular expression intended to protect against directory traversal reduces ".../...//" to "../".

**•** CVE-2005-0603: Malformed regexp syntax leads to information exposure in error message.

**•** CVE-2005-1820: Code injection due to improper quoting of regular expression.

**•** CVE-2005-3153: Null byte bypasses PHP regexp check.

**•** CVE-2005-4155: Null byte bypasses PHP regexp check.

## Related Attack Patterns (CAPEC)

* CAPEC-15
* CAPEC-6
* CAPEC-79

## Modes of Introduction

**•** Implementation: N/A

## Common Consequences

**•** Impact: Unexpected State, Varies by Context — Notes: When the regular expression is not correctly specified, data might have a different format or type than the rest of the program expects, producing resultant weaknesses or errors.

**•** Impact: Bypass Protection Mechanism — Notes: In PHP, regular expression checks can sometimes be bypassed with a null byte, leading to any number of weaknesses.

## Potential Mitigations

**•** Architecture and Design: Regular expressions can become error prone when defining a complex language even for those experienced in writing grammars. Determine if several smaller regular expressions simplify one large regular expression. Also, subject the regular expression to thorough testing techniques such as equivalence partitioning, boundary value analysis, and robustness. After testing and a reasonable confidence level is achieved, a regular expression may not be foolproof. If an exploit is allowed to slip through, then record the exploit and refactor the regular expression. (Effectiveness: N/A)

## Applicable Platforms

**•** None (Class: Not Language-Specific, Prevalence: Undetermined)

## Demonstrative Examples

**•** An attacker could provide an argument such as: "; ls -l ; echo 123-456" This would pass the check, since "123-456" is sufficient to match the "\d+-\d+" portion of the regular expression.

**•** Since the regular expression does not have anchors (CWE-777), i.e. is unbounded without ^ or $ characters, then prepending a 0 or 0x to the beginning of the IP address will still result in a matched regex pattern. Since the ping command supports octal and hex prepended IP addresses, it will use the unexpectedly valid IP address (CWE-1389). For example, "0x63.63.63.63" would be considered equivalent to "99.63.63.63". As a result, the attacker could potentially ping systems that the attacker cannot reach directly.

## Notes

**•** Relationship: While there is some overlap with allowlist/denylist problems, this entry is intended to deal with incorrectly written regular expressions, regardless of their intended use. Not every regular expression is intended for use as an allowlist or denylist. In addition, allowlists and denylists can be implemented using other mechanisms besides regular expressions.

**•** Research Gap: Regexp errors are likely a primary factor in many MFVs, especially those that require multiple manipulations to exploit. However, they are rarely diagnosed at this level of detail.