

# Flexible and Optimal Dependency Management via Max-SMT

Donald Pinckney  
Northeastern University  
Boston, USA  
pinckney.d@northeastern.edu

Federico Cassano  
Northeastern University  
Boston, USA  
cassano.f@northeastern.edu

Arjun Guha  
Northeastern University  
Boston, USA  
a.guha@northeastern.edu

Jonathan Bell  
Northeastern University  
Boston, USA  
j.bell@northeastern.edu

Massimiliano Culpo  
*np-complete, S.r.l.*  
Mantova, Italy  
massimiliano.culpo@gmail.com

Todd Gamblin  
*Lawrence Livermore National Laboratory*  
Livermore, USA  
tgamblin@llnl.gov

**Abstract**— Package managers such as NPM have become essential for software development. The NPM repository hosts over 2 million packages and serves over 43 billion downloads every week. Unfortunately, the NPM dependency solver has several shortcomings. 1) NPM is greedy and often fails to install the newest versions of dependencies; 2) NPM’s algorithm leads to duplicated dependencies and bloated code, which is particularly bad for web applications that need to minimize code size; 3) NPM’s vulnerability fixing algorithm is also greedy, and can even introduce new vulnerabilities; and 4) NPM’s ability to duplicate dependencies can break stateful frameworks and requires a lot of care to workaround. Although existing tools try to address these problems they are either brittle, rely on post hoc changes to the dependency tree, do not guarantee optimality, or are not composable.

We present PACSOLVE, a unifying framework and implementation for dependency solving which allows for customizable constraints and optimization goals. We use PACSOLVE to build MAXNPM, a complete, drop-in replacement for NPM, which empowers developers to combine multiple objectives when installing dependencies. We evaluate MAXNPM with a large sample of packages from the NPM ecosystem and show that it can: 1) reduce more vulnerabilities in dependencies than NPM’s auditing tool in 33% of cases; 2) chooses newer dependencies than NPM in 14% of cases; and 3) chooses fewer dependencies than NPM in 21% of cases. All our code and data is open and available.

**Index Terms**—package-management, Max-SMT, NPM, Rosette, dependency-management, JavaScript

## I. INTRODUCTION

Package managers such as NPM (the de facto package manager for JavaScript) have become essential for software development. For example, the NPM repository hosts over two million packages and serves over 43 billion downloads weekly. The core of a package manager is its *dependency solver*, and NPM’s solver tries to quickly find dependencies that are recent and satisfy all version constraints. Unfortunately, NPM uses a greedy algorithm that can duplicate dependencies and can even fail to find the most recent versions of dependencies.

This work is partially supported by the National Science Foundation grants CCF-2102288, CCF-2100037 and CNS-2100015.

Moreover, the users of NPM may have other goals that NPM does not serve. 1) Web developers care about minimizing code size to reduce page load times. “Bundlers” such as Webpack alter the packages selected by NPM to eliminate duplicates (Section II-B). 2) Many developers want to avoid vulnerable dependencies and several tools detect and update vulnerable dependencies, including NPM’s built-in “audit” command [1]. However, the audit command is also greedy and its fixes can introduce more severe vulnerabilities (Section II-A). 3) Finally, there are semantic reasons why many packages, such as frameworks with internal state, should never have multiple versions installed simultaneously. However, NPM’s approach to solving this, known as “peer dependencies” is brittle and causes confusion (Section II-C).

The problem with these approaches is that they are ad hoc attempts to customize and workaround NPM’s solver. Bundlers and audit tools modify solved dependencies after NPM produces its solution. Peer dependencies effectively disable the solver in certain cases and rely on the developer to select unsolved dependencies at the package root. In general, NPM cannot produce a “one-size-fits-all” solution that satisfies the variety of goals that developers have. Moreover, any tool that modifies the solver’s solution after solving risks introducing other problems and may not compose with other tools.

Our key insight is that all these problems can be framed as instances of a more general problem: optimal dependency solving, where the choices of optimization objectives and constraints determine which goals are prioritized. Due to the wide-ranging goals in the NPM ecosystem, this paper argues that NPM should allow developers to *customize and combine several objectives*. For example, a developer should be able to specify policies such as “dependencies must not have any critical vulnerabilities”, “packages should not be duplicated”, and combine these with the basic objective of “select the latest package versions that satisfy all constraints.” To make this possible and evaluate its effectiveness, we present MAXNPM: a complete, drop-in replacement for NPM, which empowers developers to combine multiple objectives.

The heart of MAXNPM is a generalized model of dependency solvers that we call PACSOLVE. PACSOLVE has a high-level DSL for specifying the syntax and semantics of package versions, version constraints, optimization objectives, and more. Under the hood, PACSOLVE uses a solver-aided language to produce a problem for a Max-SMT solver, which ensures that its solution is optimal, unlike NPM’s greedy approach. Although we apply PACSOLVE to build a package manager for NPM, we believe that the generality of PACSOLVE will make it possible to build customizable dependency solvers for other package ecosystems as well.

We use MAXNPM to conduct an empirical evaluation with a large dataset of widely-used packages from the NPM repository. Our evaluation shows that MAXNPM outperforms NPM in several ways:

- 1) chooses newer dependencies compared to using NPM for 14% of packages with at least one dependency.
- 2) shrinks the footprint of 21% of packages with at least one dependency.
- 3) reduces the number or severity of security vulnerabilities in 33% of packages with at least one dependency.

Overall, MAXNPM takes just 2.6s longer than NPM on average, though encounters some outliers which solve significantly more slowly with MAXNPM, which is reflected in the standard deviation of the slowdown (13.7s).

This paper makes the following contributions:

- 1) PACSOLVE: an executable semantics for dependency solvers, which are the key component of package managers. PACSOLVE is parameterized along several key axes, which allows for customization of constraints and optimization objectives.
- 2) MAXNPM: a drop-in replacement for NPM that allows the user to solve dependencies with several objectives, including maximizing secure dependencies, decreasing code size, and maximizing up-to-date dependencies. MAXNPM is implemented by instantiating PACSOLVE to use NPM’s notions of versions and version constraints.

## II. BACKGROUND

NPM is the most widely used package manager for JavaScript. (Alternatives such as Yarn are compatible with NPM configurations.) NPM is co-designed with Node, which is a popular JavaScript runtime for desktop and server applications. However, NPM is also widely used to manage web applications’ dependencies, using “bundlers” like Webpack to build programs for the browser.

An NPM configuration (the `package.json` file) lists package dependencies with version constraints. NPM has a rich syntax for version constraints that MAXNPM fully supports. However, a typical configuration specifies version ranges or exact versions of each dependency.

An unusual feature of NPM is that it may select multiple versions of the same package. To illustrate, consider the following real-world example that involves the packages `debug` and `ms` (calendar utilities). Suppose a project depends

on the latest version of `ms` (version 2.1.3) and `debug`. Unfortunately, `debug` depends on an older version of `ms` (version 2.1.2). In this situation, NPM selects both versions of `ms`: the root of the project get the latest version, and `debug` gets the older version, thus satisfying both constraints independently instead of failing to unify the two constraints. Some package managers handle this situation differently: PIP would report that no solution exists, while Maven and NuGet would install only the newer version of `ms` and let `debug` load a different version than what it asked for.

Unfortunately, NPM’s behavior is not always desirable, and can lead to increased code size and subtle runtime bugs. Moreover, NPM does not guarantee that packages are only duplicated when strictly necessary.

Another problem is that what it means to be a *newer package* is not well-defined across package managers. Suppose package *A* has versions 1.0.0 and 2.0.0, and then its author publishes a security numbered 1.0.1. What happens if a program depends on *A* with no constraints? Surprisingly, NPM will choose 1.0.1 because it was uploaded last, while some other package managers (such as PIP and Cargo) would select 2.0.0. Without specific knowledge of the situation, it is unclear which choice is better.

### A. Avoiding Vulnerable Dependencies

NPM’s built-in tool `npm audit` checks for vulnerable dependencies by querying the GitHub Advisory Database. The tool can also fix vulnerabilities by upgrading dependencies without violating version constraints.<sup>1</sup> However, the tool has several shortcomings. 1) Each run only tries to fix a single vulnerability. 2) It only upgrades vulnerable dependencies, even if a vulnerability-free downgrade is available that respects version constraints. 3) It does not prioritize fixes by vulnerability scores (CVSS), even though they are available in the GitHub Advisory Database. 4) It does not make severity-based compromises. For example, a fix may introduce new vulnerabilities that are more severe than the original.

### B. Minimizing Code Bloat

A “bundler”, such as Webpack, Browserify, or Parcel, is a tool that works in concert with NPM to manage the dependencies of front-end web applications. The primary task of a bundler is to package all dependencies to be loaded over the web, instead of the local filesystem. However, bundlers do more, including work to minimize page load times. A simple way to minimize page load times is to reduce code size. Unfortunately, NPM’s willingness to duplicate packages can lead to increased code size [2]. Contemporary bundlers employ a variety of techniques from minification to unifying individual files with identical contents. However, these techniques are not always sound and have been known to break widely-used front-end frameworks [3], [4].

<sup>1</sup>The `--force` flag breaks constraints and potentially breaks the program.

### C. Managing Stateful Dependencies

NPM’s ability to select several versions of the same dependency is also unhelpful when using certain stateful frameworks. For example, React is a popular web framework that relies on internal global state to schedule view updates. If a program depends on two packages that transitively depend on two different versions of React, it is likely to encounter runtime errors or silent failures. The only way to avoid this problem is if all package authors are careful to mark their dependency on React as a *peer dependency*: a dependency that is installed by some other package in a project. However, there is no easy way to determine that all third-party dependencies use peer dependencies correctly. It can also be hard to determine before hand that a package will never be used as a dependency and thus should not use peer dependencies.

### III. MAXNPM

The goal of MAXNPM is to help developers address the broad range of problems described above. MAXNPM serves as a drop-in replacement for the default `npm install` command. The user can run `npm install --maxnpm` to use MAXNPM’s customizable dependency solver instead. There are two broad ways to customize MAXNPM. First, MAXNPM allows the user to specify a prioritized list of objectives with the `--minimize` flag. Out of the box, MAXNPM supports the following objectives (defined precisely in Figure 3):

- `min_oldness`: minimizes the number and severity of installed old versions;
- `min_num_deps`: minimizes the number of installed dependencies;
- `min_duplicates`: minimizes the number of co-installed different versions of the same package; and
- `min_cve`: minimizes the number and severity of known vulnerabilities.

Second, MAXNPM allows the user to customize how multiple package versions are handled with the `--consistency` flag:

- `npm`: the default behavior of NPM, which allows several versions of a package to co-exist in a single project; and
- `no-dups`: require every package to have only one version installed.

With some work, the user can even define new objectives and consistency criteria.

For example, a developer building a front-end web application may want to reduce code size and select recent package versions. They could use MAXNPM as follows:

---

```
1 npm install --maxnpm --consistency no-dups
2   --minimize min_oldness,min_num_deps
```

---

This command avoids duplicating dependencies, and minimizes oldness and the number of dependencies in that order. This is a more principled approach to reducing code size than the ad hoc de-duplication techniques used by bundlers. Moreover, we show that this command is frequently successful at reducing code size (Section IV-A3).

#### Package Metadata

$\mathcal{P} ::= \text{String}$	Package Names
$\mathcal{V}$ is a set	Version Numbers
$\mathcal{C}$ is a set	Version Constraints
$\mathcal{D} ::= \mathcal{P} \times \mathcal{C}$	Dependencies
$\mathcal{N}' \subseteq \mathcal{P} \times \mathcal{V}$	Package repository (finite set)
$\mathcal{N} ::= \mathcal{N}' \cup \{\text{root}\}$	The root node of the solve
$\text{deps} : \mathcal{N} \xrightarrow{\text{fin}} \mathcal{D}^*$	Dependencies per node
$\mathcal{M} ::= \langle \mathcal{N}, \text{deps} \rangle$	Package metadata

#### Solution Graph

$N_R \subseteq \mathcal{N}$	Package versions in solution
$D_R \in N_R \rightarrow N_R^*$	Solved dependencies
$\mathcal{G} ::= \langle N_R, D_R \rangle$	Solution graphs

#### Dependency Solver Specification

$\text{sat} : \mathcal{C} \rightarrow \mathcal{V} \rightarrow \text{Bool}$	Constraint satisfaction semantics
$\text{consistent} : \mathcal{V} \rightarrow \mathcal{V} \rightarrow \text{Bool}$	Version consistency versions
$\text{minGoal} : \mathcal{G} \rightarrow \mathbb{R}^n$	Objective functions

Fig. 1: The PACSOLVE Model of Dependency Solving

As a second example, consider a developer building a web application backend, where they are very concerned about security vulnerabilities. They could use MAXNPM as follows:

---

```
npm install --maxnpm --minimize min_cve,min_oldness
```

---

This command subsumes `npm audit fix` and we show that it is substantially more effective (Section IV-A1).

MAXNPM is built on PACSOLVE, which is a DSL for describing dependency solvers. This section presents PACSOLVE with an emphasis on how we use it to build MAXNPM. Section VI discusses the potential applicability of PACSOLVE to building solvers for other languages.

#### A. Describing a Dependency Solver with PACSOLVE

PACSOLVE is a solver-aided language built with Rosette, a solver-aided programming language that interfaces with the Z3 theorem prover [5]. The input to PACSOLVE is a set of available packages with their dependencies. The output is a *solution graph* where nodes are specific versions of a package and edges represent dependencies. (Section III-B formally describes the semantics and relationship between available packages and solution graphs.) Without loss of generality, we assume there is a distinguished root package (`root`).

To build a dependency solver with PACSOLVE, we have to define:

- 1) The abstract syntax of package versions ( $\mathcal{V}$ ) and version constraints ( $\mathcal{C}$ ), which tends to vary subtly between dependency solvers;
- 2) The *constraint satisfaction predicate* that determines if a given version satisfies a constraint (*sat*);
- 3) The *version consistency predicate* that consumes two package versions and determines if those two versions of the same package may be co-installed (*consistent*); and
- 4) The *objective function* that determines the cost of a solution graph (*minGoal*).

$\mathcal{V} := (x\ y\ z)$	Version numbers
$\mathcal{C} := (= x\ y\ z)$	Exact
*	Any
( $\leq x\ y\ z$ )	At most
( $\geq x\ y\ z$ )	At least
( $\wedge x\ y\ z$ )	Semver compatible with
( $\text{and } \mathcal{C}_1\ \mathcal{C}_2$ )	Conjunction
( $\text{or } \mathcal{C}_1\ \mathcal{C}_2$ )	Disjunction

(a) Example of  $\mathcal{V}$  and  $\mathcal{C}$ , allowing conjunction, disjunction, and range operators on semver-style versions.

```

1 (define (sat c v)
2   (match ` (,v ,c)
3     [ `((,x ,y ,z) (= ,x ,y ,z))      #true]
4     [ `(_ ,*)                          #true]
5     [ `((,x ,y ,z1) (<= ,x ,y ,z2))   (<= z1 z2)]
6     [ `((,x ,y1 ,z1) (<= ,x ,y2 ,z2)) (< y1 y2)]
7     [ `((,x1 ,y1 ,z1) (<= ,x2 ,y2 ,z2)) (< x1 x2)]
8     [ `((0 0 ,z1) (^ 0 0 ,z2))        (= z1 z2)]
9     [ `((0 ,y ,z1) (^ 0 ,y ,z2))      (>= z1 z2)]
10    [ `((0 ,y1 ,z1) (^ 0 ,y2 ,z2))     #false]
11    [ `((,x ,y ,z1) (^ ,x ,y ,z2))    (>= z1 z2)]
12    [ `((,x ,y1 ,z1) (^ ,x ,y2 ,z2))  (> y1 y2)]
13    [ `(_ (and ,c1 ,c2))
14      (and (sat c1 v) (sat c2 v))]
15    [ `(_ (or ,c1 ,c2))
16      (or (sat c1 v) (sat c2 v))]
17    [ `((,x1 ,y1 ,z1) (>= ,x2 ,y2 ,z2))
18      (sat ` (,x2 ,y2 ,z2) ` (<= ,x1 ,y1 ,z1))]
19    [ _                                  #false])
20
21 (define (consistent v1 v2)
22   #true)

```

(b) A constraint satisfaction predicate and a version consistency predicate for  $\mathcal{V}$  and  $\mathcal{C}$  defined in Figure 2a.

Fig. 2: The syntax of versions and constraints, and the constraint satisfaction predicate and consistency predicate for a fragment of NPM.

PACSOLVE produces a solution that is optimal and consistent with all constraints. A solution graph may optionally have cycles, which some package managers allow (including NPM).

*A Fragment of NPM:* Figure 2a shows an example of versions and constraints for a fragment of NPM. (MAXNPM supports the full syntax and takes care of parsing NPM’s concrete syntax to the parenthesized syntax that Rosette and PACSOLVE require.) Given this syntax of versions and constraints, Figure 2b shows constraint satisfaction and version consistency predicates. The `sat` function receives as input a version constraint and a version (from the syntax of Figure 2a) and returns a boolean by performing a `match` case analysis. Interesting subtleties include lines 8–12 handling the matching semantics of caret constraints (`^1.2.3`) including their different behavior for versions with leading zeros, and lines 13–16 performing structural recursion on the constraints. The `consistent` function receives two versions and must determine if they can be co-installed. This simple consistency predicate is the constant `true` function, which implements NPM’s standard policy of always allowing co-installations.

Figure 3 defines three examples of objective functions that consume a solution graph and produce a cost. PACSOLVE

minimizes cost, so a trivial objective is to minimize the total number of packages, completely ignoring package age and other factors (Figure 3a). An alternative objective is to minimize the co-installation of multiple versions of the same package (Figure 3b). The function counts the number of versions of each package, and assigns a cost to every package that has more than one version. Our final example is an interpretation of the common goal that package managers have of trying to choose newer versions of dependencies (Figure 3c). The function gives each node an *oldness score* between 0 (newest) and 1 (oldest), with the scores evenly divided across all versions of a package in the solution. There are two subtleties with this definition. 1) We perform *minimization*, since maximization would encourage the solver to find large solutions that inflate newness. 2) We take the sum rather than the mean, since taking the mean would also encourage the solver to add extra packages that deflate oldness.

Other metrics are possible as well, such as our implementation of an aggregated score of dependency vulnerabilities (Section IV-A1), or the total download size [6]. With a library of optimization objective functions defined, PACSOLVE then allows for easy composition of objective functions either by multiple prioritized objectives, or by weighted linear combinations of objectives.

## B. The Semantics of PACSOLVE

This section formally describes the semantics of PACSOLVE. The *package metadata* (top of Figure 1) is described by 1) a set of package name and version pairs ( $\mathcal{N}$ ), and 2) a map (*deps*) from these pairs to a list of dependencies. Each dependency specifies a package name and version constraint ( $\mathcal{C}$ ).

The solution graph is a directed graph where the nodes are package-version pairs, and each node has an ordered list of edges. The order of edges corresponds to the order of dependencies in the package metadata.

The semantics of PACSOLVE is a relation ( $\mathcal{S}$ ) between the package metadata, the dependency solver specification, and the solution graph. The relation holds when a solution graph is valid with respect to the package metadata and the dependency solver specification. The relation holds if and only if the following six conditions are satisfied. First, the solution graph must include the root:

$$\text{root} \in N_R \quad (1)$$

Second, the solution graph must be connected, to ensure it does not have extraneous packages:

$$\langle N_R, D_R \rangle \text{ is connected} \quad (2)$$

Third, for all packages in the solution graph, every edge must correspond to a constraint in the package metadata:

$$\forall n.n \in N_R \implies |D_R(n)| = |\text{deps}(n)| \quad (3)$$

Fourth, for every edge in the solution graph that points to package  $p$  with version  $v$ , the corresponding constraint in the

```

1 (define (minGoal-num-deps g)
2   (length (graph-nodes g)))

```

(a) Minimize the total number of installed dependencies.

```

1 (define (minGoal-duplicates g)
2   ; we count how many times
3   ; each package name occurs
4   (define package-counts (foldl
5     (lambda (n counts)
6       (define p (node-package n))
7       (hash-set counts p
8         (add1 (hash-ref counts p 0))))
9     (make-immutable-hash)
10    (graph-nodes g)))
11
12 ; then assign a cost of 1
13 ; for each duplicate
14 (apply +
15   (map
16     (lambda (c) (max 0 (sub1 c)))
17     (hash-values package-counts)))

```

(b) Minimize the total number of co-installed versions of the same package.

```

1 (define (minGoal-oldness g)
2   (apply +
3     (map
4       (lambda (n)
5         (get-oldness
6           (node-package n)
7           (node-version n)))
8       (graph-nodes g))))
9
10 (define (get-oldness p v)
11 ; The get-sorted-versions retrieves
12 ; a list of all versions of p
13 (define all-vs
14   (get-sorted-versions p))
15 (if (= (length all-vs) 1)
16     0
17     (/ (index-of all-vs v)
18        (sub1 (length all-vs)))))

```

(c) Minimize the amount of “oldness” present in the solution graph. Each resolved dependency contributes an oldness proportional to its rank among the total ordering of versions of that package

Fig. 3: Three different examples of PACSOLVE minimization objectives

package metadata must refer to package  $p$  with constraint  $c$ , where  $v$  satisfies  $c$ :

$$\begin{aligned}
& \forall n. n \in N_R \implies \\
& \quad \forall i. 0 \leq i < |D_R(n)| \implies \\
& \exists p, v, c. (p, v) = D_R(n)[i] \\
& \quad \wedge (p, c) = \text{deps}(n)[i] \wedge \text{sat}(c, v)
\end{aligned} \tag{4}$$

The criteria so far are adequate for many dependency solvers, but permits solutions that may be unacceptable. For example, without further constraints, a solution graph may have several versions of the same package. Thus the fifth condition ensures that if there are versions of a package in the solution graph, then the two versions are consistent, as judged by the dependency solver specification:

$$\forall p, v, v'. (p, v), (p, v') \in N_R \implies \text{consistent}(v, v') \tag{5}$$

NPM allows arbitrary versions to be co-installed (so *consistent* is the constant true function), PIP only allows exactly one version of a package to be installed at a time (so *consistent* requires  $v = v'$ ), and Cargo only allows semver-incompatible versions to be co-installed.

A final distinction between dependency solvers is whether or not they allow cyclic dependencies. NPM and PIP allow cycles, but others, such as Cargo, do not. Thus the sixth condition, which is *optional*, uses the dependency solver specification to determine whether or not cycles are permitted:

$$\langle N_R, D_R \rangle \text{ is acyclic} \tag{6}$$

These five or six conditions determine whether or not a solution graph is correct with respect to the semantics of a particular dependency solver.

### C. Synthesizing Solution Graphs with PACSOLVE

Dependency solving with possible conflicts is NP-complete [7]. Some package managers use polynomial-time

algorithms by giving up on various properties, such as disregarding conflicts (NPM) and eschewing completeness (NPM and PIP’s old solver [8]). Since the PACSOLVE model includes a generalized notion of conflicts (*consistent*), we leverage Max-SMT solvers to implement PACSOLVE effectively.

We implement PACSOLVE in Rosette, which is a solver-aided language that facilitates building verification and synthesis tools for DSLs. In the PACSOLVE DSL, *the program is a solution graph*. We implement a function that consumes 1) package metadata, 2) a dependency solver specification (Figure 1) and 3) a solution graph, and then asserts that the solution graph is correct. With a little effort, we can replace the input solution graph with a *solution graph sketch*. This allows us to use Rosette to perform *angelic execution* [9] to synthesize a solution graph that satisfies the correctness criteria. This section describes the synthesis procedure in more detail, starting with how we build a sketch.

*a) Sketching solution graphs:* Before invoking the Rosette solver, we build a sketch of a solution graph that has a node for every version of every package that is reachable from the set of root dependencies. Every node in a sketch has the following fields: 1) a concrete name and version for the package that it represents; 2) a symbolic boolean *included* that indicates whether or not the node is included in the solution graph; 3) a symbolic natural number *depth* which we use to enforce acyclic solutions when desired; 4) a vector of concrete dependency package names; 5) a vector of concrete version constraints for each dependency; and 6) a vector of symbolic *resolved versions* for each dependency.

Figure 4 illustrates an example solution graph sketch corresponding to a dependency solving problem involving two packages (*debug*, *ms*), where *ms* is a dependency of both *debug* and the root, while *debug* is depended on by only the root. The combination of concrete dependency names and

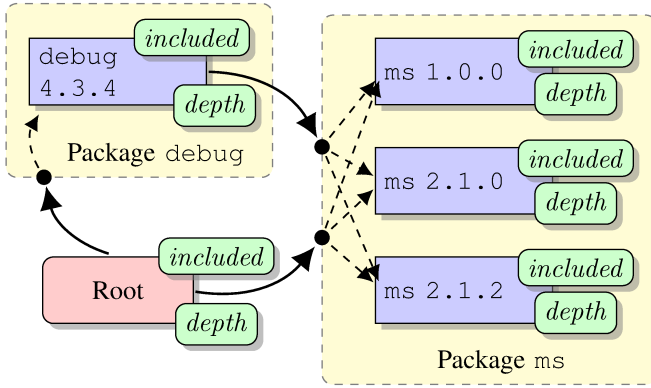


Fig. 4: A sample solution graph sketch

symbolic dependency versions can be seen as representing symbolic edges in two parts: a concrete part which does not need to be solved for (solid arrows), and a symbolic part which requires solving (dashed arrows). This representation shrinks the solution space of graphs as outgoing edges can only point to nodes with the correct package name.

b) *Graph Sketch Solving*: We define three assertion functions that check correctness criteria of a solution graph:

- 1) `check-dependencies` asserts that if a node is *included*, then all the dependencies of the node are *included* and satisfy the associated version constraints as judged by the constraint interpretation function (*sat*). We run this assertion function on all nodes, and additionally assert that the root node is *included*.
- 2) `globally-consistent?` asserts that the consistency function (*consistent*) holds on all pairs of nodes with the same package name.
- 3) `acyclic?` asserts that the *depth* of a node is strictly less than the depth of all its dependencies. If an acyclic solution is desired, we run this assertion function on all nodes, and assert that the root node has depth zero.

We then ask Rosette to find a concrete solution graph that satisfies the above constraints while minimizing the objective function (*minGoal*). As a final step, we traverse the concretized solution graph sketch from the root node, and install on-disk those nodes that are marked *included*.

#### IV. EVALUATION

We evaluate MAXNPM in several scenarios, determining whether its support for flexible optimization objectives can provide tangible benefits to developers as compared to NPM. We gather two large datasets of popular packages, and also investigate if MAXNPM is sufficiently reliable and performant to use as a drop-in replacement for NPM. In this section we use MAXNPM to answer each of our research questions:

**RQ1:** Can MAXNPM find better solutions than NPM when given different optimization objectives?

**RQ2:** Do MAXNPM’s solutions pass existing test suites?

**RQ3:** Does MAXNPM successfully solve packages that NPM solves?

**RQ4:** Does using MAXNPM substantially increase solving time?

We build two datasets of NPM packages. *Top1000* is a set of the latest versions of the top 1,000 most-downloaded packages as of August 2021. Including their dependencies, there are 1,147 packages in this set. Unsurprisingly, these packages are maintained and have few known vulnerabilities. Therefore, to evaluate vulnerability mitigation, we build the *Vuln715* dataset of 715 packages with high CVSS scores as follows: 1) we filter the *Top1000* to only include packages with available GitHub repositories; 2) we extract every revision of `package.json`; 3) for each revision, we calculate the aggregate CVSS score of their direct dependencies, as determined by the GitHub Advisory Database; and 4) we select the highest scoring revision of each package.

MAXNPM is built on NPM 7.20.1. PACSOLVE uses Racket 8.2, Rosette commit `1d042d1`, and Z3 commit `05ec77c`. We configure NPM to not run post-install scripts and not install optional dependencies. We run our performance benchmarks on Linux, with a 16-Core AMD EPYC 7282 CPU with 64 GB RAM. We warm the NPM local package cache before measuring running times.

A. *RQ1: Can MAXNPM find better solutions than NPM when given different optimization objectives?*

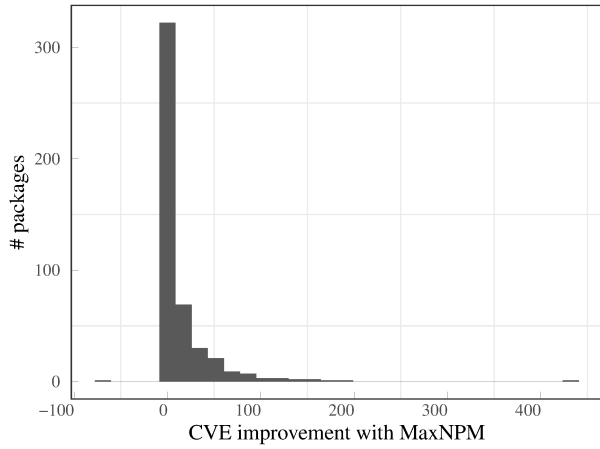
1) *Can MAXNPM help avoid vulnerable dependencies?:* We configure MAXNPM to minimize the aggregate CVSS scores of all dependencies,<sup>2</sup> and compare with the built-in `npm audit fix` tool (Section II-A). We use the *Vuln715* packages for this comparison. Both tools run successfully on 472 packages: the failures occur because these are typically older versions that do not successfully install.

The histogram in Figure 5a reports the difference in aggregate CVSS score between `npm audit fix` and MAXNPM. A higher score indicates that a package has fewer vulnerabilities with MAXNPM. MAXNPM produces fewer vulnerabilities on 235 packages (33%). There is one package where MAXNPM produces a lower score; we are investigating this as a possible bug. The mean CVSS improvement by MAXNPM is 14.75 CVSS points (a “maximum severity” vulnerability is 10 points), or by 30.51%. The improvement is statistically significant ( $p < 2.2 \times 10^{-16}$ ) using a paired Wilcoxon signed rank test, with a medium Cohen’s *d* effect size of  $d = 0.46$ . Thus, we find that MAXNPM is substantially more effective than `npm audit fix` at removing vulnerable dependencies.

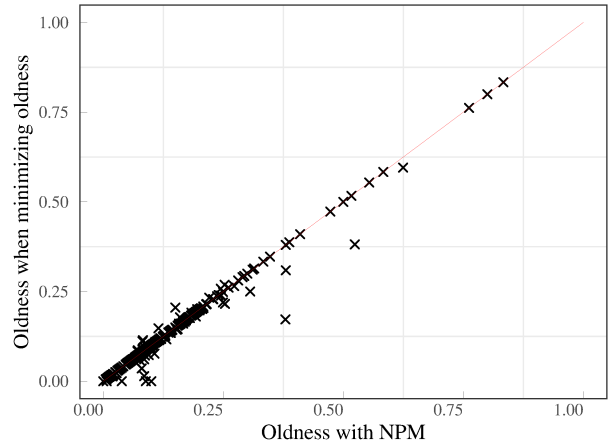
An example project where MAXNPM eliminates vulnerabilities is the `babel` compiler (34 million weekly downloads). Commit `5b09114b8` is in *Vuln715*, and MAXNPM eliminates all vulnerabilities; whereas `npm audit fix` leaves several with an aggregate CVSS score of 59.4.

2) *Can MAXNPM find newer packages than NPM?:* MAXNPM ought to be able to find newer packages than NPM’s greedy algorithm. We define the oldness of a dependency on a package version ( $old(p, v)$ ) as a function that

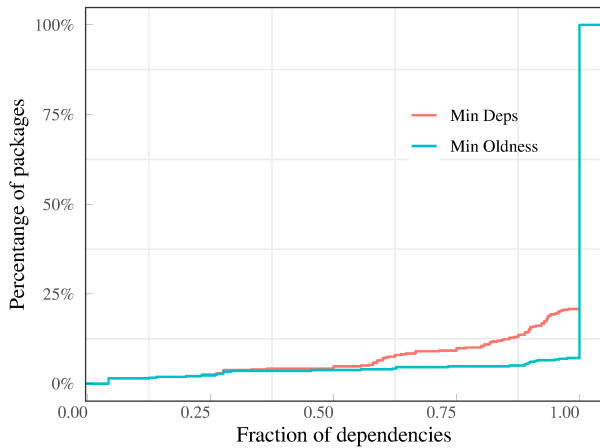
<sup>2</sup>The `min_cve`, `min_oldness` flags.



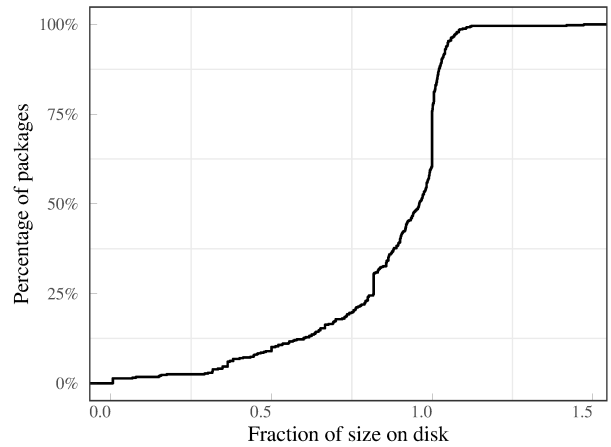
(a) A histogram showing CVSS improvement of packages solved with MAXNPM (configured to minimize vulnerabilities first) compared to NPM’s auditing tool. CVSS decreases with 33% of packages.



(b) MAXNPM is configured to minimize oldness and the number of dependencies (in that order). Each point represents a package, and those below the line have newer dependencies, by the metric in Section IV-A2. Overall, MAXNPM finds newer versions for dependencies.



(c) An ECDF of the fraction of dependencies compared to NPM, when MAXNPM is configured to a) minimize dependencies then oldness (red ECDF), or b) minimize oldness then dependencies (blue ECDF). MAXNPM can reduce dependencies in about 21% of packages.



(d) An ECDF of the ratio of disk space of packages solved using MAXNPM (configured to minimize number of dependencies, then oldness) vs NPM. MAXNPM can reduce space required.

Fig. 5: Comparing NPM’s to MAXNPM’s solution quality. These plots ignore failures in both solvers and have MAXNPM configured to use NPM-style consistency and allow cycles.

assigns the newest version the value 0, the oldest version the value 1, and other versions on a linear scale in between. We define the mean oldness of project as the mean oldness of all dependencies in a project, including transitive dependencies. Note that this metric is not identical to the minimization objective of MAXNPM, which calculates the sum and ignores duplicates. The metric is more natural to interpret, whereas the objective function avoids pathological solutions.

Figure 5b shows a point for every package in the *Top1000*, with mean oldness using NPM and MAXNPM as its  $x$  and  $y$  coordinates. Points on  $y = x$  are packages whose dependencies are just as old with both NPM and MAXNPM.

14% of packages excluding those with zero dependencies are better with MAXNPM, while 5% are worse. On average oldness improved by 2.62%. The improvement is statistically significant ( $p = 4.27 \times 10^{-6}$ ) using a paired Wilcoxon signed rank test, with a small Cohen’s  $d$  effect size of  $d = 0.024$ . Thus MAXNPM produces newer dependencies on average.

An example of successful oldness minimization is the `class-utils` package (15 million weekly downloads). MAXNPM chooses a slightly older version of a direct dependency, which allows it to chose much newer versions of transitive dependencies.

One might wonder why MAXNPM does worse in 5% of

cases, since MAXNPM should be optimal. Manual investigation of these cases shows that some packages make use of features which we have not implemented in MAXNPM, such as URLs to tarballs rather than named dependencies. MAXNPM is unable to explore that region of the search space. Implementing these features would take some engineering effort, but wouldn't require changes to the model.

3) *Can MAXNPM reduce code bloat?*: Instead of using ad hoc and potentially unsound techniques to reduce code bloat, we can configure MAXNPM to minimize the total number of dependencies. On the *Top1000* packages, we configure MAXNPM in two ways: 1) prioritize fewer dependencies over lower oldness; and 2) prioritize lower oldness over fewer dependencies. Figure 5c plots an ECDF (empirical cumulative distribution function) plot where the  $x$ -axis shows the shrinkage in dependencies compared to NPM, and the  $y$ -axis shows the cumulative percentage of packages with that amount of shrinkage. The plot excludes packages with zero dependencies, and  $x = 1$  indicates no shrinkage (when MAXNPM produces just as many dependencies as NPM).

Both configurations produce fewer dependencies than NPM, but prioritizing fewer dependencies is the most effective (the red line). For about 21% of packages, MAXNPM is able to reduce the number of dependencies, with an average reduction in the number of packages of 4.37%. The improvement is statistically significant ( $p < 2.2 \times 10^{-16}$ ) using a paired Wilcoxon signed rank test, with a moderate Cohen's  $d$  effect size of  $d = 0.20$ . For the same set of packages, the total disk space required shrinks significantly (Figure 5d). With MAXNPM, a quarter of the packages require 82% of their original disk space. Even when we prioritize lowering oldness, MAXNPM still produces fewer dependencies (the blue line).

An example of dependency size minimization is the `assert` package (13 million weekly downloads). For 3 direct dependencies, MAXNPM chooses slightly older revisions (with the same major and minor version). But this eliminates 33 of 43 transitive dependencies.

We have observed that in a few cases NPM exhibits a bug in which it installs additional dependencies that are not defined in the set of production dependencies of the package, nor in the set defined by transitive dependencies.<sup>3</sup> However, MAXNPM does not exhibit this bug. We will work on reporting this bug, but we believe this an example of the advantage of PACSOLVE's declarative style of building package managers.

4) *Can MAXNPM address duplicated packages?*: NPM happily allows a program to load several versions of the same package, which can lead to subtle bugs (Section II-C). To address this problem, a developer can configure MAXNPM to disallow duplicates. In this configuration, 19 packages (1.9%) in *Top1000* produce unsatisfiable constraints, which indicates that they require several versions of some package (Table I).

For example, `terser@5.9.0` is a widely used JavaScript parser that that directly depends on `source-map@0.7.x`

<sup>3</sup>`@babel/plugin-proposal-export-namespace-from` is an example.

MAXNPM	Pass	563 77%	7 1%
	Fail	38 5%	127 17%
		Pass	Fail
		NPM	

Fig. 6: Results of running tests after solving dependencies with NPM and MAXNPM. In total only 5% of packages have a failing test with MAXNPM but not with NPM.

and `source-map-support@0.5.y`. However, the latter depends on `source-map@0.6.z`, thus the build must include both versions of `source-map`. The ideal fix would update `source-map-support` to support `source-map@0.7.x`.

#### B. RQ2: Do MAXNPM's solutions pass existing test suites?

Although a package should pass its test suite with any set of dependencies that satisfy all constraints, in practice tests may fail with alternate solutions due to under-constrained dependencies. We identified test suites for 735 of the *Top1000* packages. A test suite succeeds only if all tests pass. All test suites succeed with both MAXNPM and NPM on 77% of packages, and fail for both on 17%. There are 38 packages where MAXNPM fails but NPM passes, and 7 packages where NPM fails and MAXNPM succeeds (Figure 6).

Test failures occurring slightly more often with MAXNPM's solutions are likely due to the fact that many of these packages have already been solved and tested with NPM, so even if their dependencies are underconstrained, at present NPM produces working solutions. Manual investigation suggests that the 7 packages that fail with NPM but succeed with MAXNPM are likely due to flaky tests and missing development dependencies while the 38 packages that fail with MAXNPM but succeed with NPM are due to those reasons in addition to under-constrained dependencies.

Finally, to verify that packages which pass their tests with MAXNPM are not doing so vacuously due to no or few tests, in Table II we report statistics of the number of executed tests for packages in the group that pass with NPM and MAXNPM, and in the group that pass with NPM and fail with MAXNPM. A two-sided Mann-Whitney U test indicates that there is no statistically significant difference between the two populations ( $p = 0.49$ ).

#### C. RQ3: Does MAXNPM successfully solve packages that NPM solves?

The rightmost three columns of Table I show the number of failures resolving dependencies on the *Top1000* for NPM, along with each configuration of MAXNPM that we evaluated.



Solver	Consistency	Allow cycles?	Minimization Objectives		Successes	Failures		
			Primary	Secondary		Unsat	Timeout	Other
NPM					953	0	0	47
MAXNPM	npm	Yes	Oldness	# of Dependencies	972	0	27	1
MAXNPM	npm	Yes	# of Dependencies	Oldness	972	0	27	1
MAXNPM	npm	Yes	Oldness	Duplicate Packages	973	0	26	1
MAXNPM	no-dups	Yes	Oldness	# of Dependencies	926	19	54	1
MAXNPM	npm	No	Oldness	# of Dependencies	972	0	27	1
MAXNPM	no-dups	No	Oldness	# of Dependencies	926	19	54	1

TABLE I: Failures that occur when running NPM and different configurations of MAXNPM on the *Top1000* dataset.

Statistic	NPM Pass & MAXNPM Pass	NPM Pass & MAXNPM Fail
Mean	489.36	58.90
STD	4699.01	266.87
Minimum	0.00	0.00
25th Perc.	0.00	1.50
50th Perc.	7.00	4.00
75th Perc.	39.00	10.50
Maximum	81582.00	1493.00

TABLE II: Statistics of the number of executed tests per package in the top left and top right groups of Figure 6.

On the *Top1000* packages, NPM itself fails on 47 packages. Many of these failures occur due to broken, optional peer-dependencies that MAXNPM does not needlessly solve.<sup>4</sup> We run MAXNPM in several configurations, and get 26–28 failures when *duplicate versions are permitted*. Some failures occur across all configurations, e.g., one package requires macOS. Most of our other failures are timeouts: we terminate Z3 after 10 minutes. When duplicates are not permitted, we do get more failures due to unsatisfiable constraints, but these are expected (Section IV-A4). Some users may prefer to have MAXNPM fail when it cannot find a solution rather than falling back to an unconstrained solution, as the latter may lead to subtle and hard-to-debug issues at runtime due to e.g. conflicting global variables in multiple versions of the same package. When we permit duplicates like NPM, we find that MAXNPM successfully builds *more* packages than NPM itself, providing strong evidence that MAXNPM can reliably be used as a drop-in replacement for NPM.

#### D. RQ4: Does using MAXNPM substantially increase solving time?

On the *Top1000* packages, we calculate how much *additional time* MAXNPM takes to solve dependencies over NPM. We observe that the minimum slowdown is  $-2.3s$  (when MAXNPM is faster than NPM), the 1st quartile is 0.8s, the median is 1.6s, the mean is 2.6s, the 3rd quartile is 2.2s, the max is 329s, and the standard deviation of the slowdown is 13.7s. These absolute slowdowns are on top of the baseline of NPM, which takes 1.52s on average, and 1.34s at the median. We exclude timeouts from this analysis, we report those in Table I. As evidenced by the maximum and

<sup>4</sup>They are not necessary to build, but NPM attempts to solve for them even with the `--omit-peer` flag.

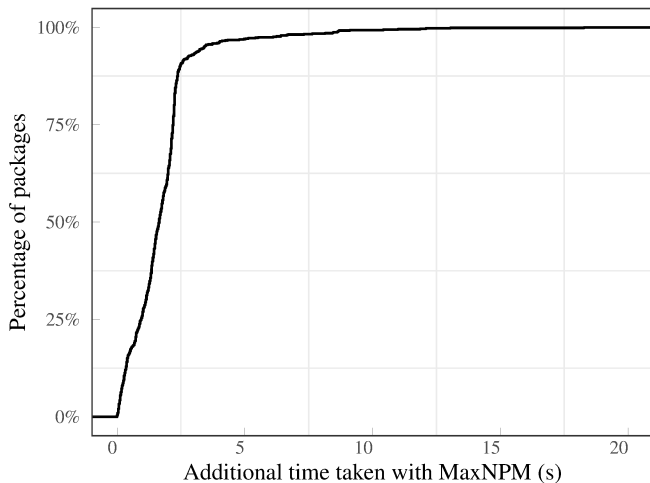


Fig. 7: ECDF of the additional time taken by MAXNPM to solve and install packages compared to NPM, ignoring timeouts and failures, with outliers ( $> 20s$ ) excluded. The outliers take up to 329s extra seconds, but the mean and median slowdowns are only 2.6s and 1.6s, respectively. In this experiment MAXNPM was configured with NPM-style consistency, allowing cycles, and minimizing oldness first and then number of dependencies.

standard deviation, there are a few outliers where MAXNPM takes substantially longer. We also perform a paired Wilcoxon signed rank test and find that the slowdown is statistically significant ( $p < 2.2 \times 10^{-16}$ ), with a moderate Cohen’s  $d$  effect size of  $d = 0.27$ . Figure 7 shows an ECDF of the absolute slowdown, but with outliers ( $> 20s$ ) removed. We conclude that while MAXNPM does increase solving time, the increase is modest in the majority of cases, but there are a few outliers. This performance characteristic mirrors that of other SAT-solver based package managers, including production ones such as Conda [10].

Excluding outliers, a significant portion of the overhead is serializing data between JavaScript (MAXNPM) and Racket (PACSOLVE), which could be improved by building the solver in JavaScript or using a more efficient serialization protocol. As for the outliers, one could implement a tool that first tries MAXNPM but reverts to greedy solving after a timeout, at the expense of optimality.

## V. RELATED WORK

Van der Hoek et al. first discussed the idea of “software release management” [11] for large numbers of independent packages in 1997, and the first package managers for Linux distributions emerged at around the same time [12], [13]. The version selection problem was first shown to be NP-complete and encoded as a SAT and Constraint Programming (CP) problem by Di Cosmo et al. [7], [14] in 2005. This early work led to the Mancoosi project, which developed the idea of a modular package manager with customizable solvers [15], [16]. This work centers around the Common Upgradeability Description Format (CUDF), an input format for front-end package managers to communicate with back-end solvers.

CUDF facilitated the development of solver *implementations* using Mixed-Integer Linear Programming, Boolean Optimization, and Answer Set Programming [17], [18], [19], and many modern Linux distributions have adopted CUDF-like approaches [20]. OPIUM [6] examined the use of ILP with weights to minimize the number of bytes downloaded or the total number of packages installed.

While package managers have their roots in Linux distributions, they have evolved considerably since the early days. Modern language ecosystems have evolved their own package managers [21], [22], [23], [24], with solver requirements distinct from those of a traditional Linux distribution. Distribution package managers typically manage only a single, global installation of each package, while language package managers are geared more towards programmers and allow *multiple* installations of the same software package.

For the most part, language ecosystems have avoided using complete solvers. As we have found in our implementation, solvers are complex and interfacing with them effectively is more challenging compared to implementing a greedy algorithm. Even on the Linux distribution side, so-called *functional* Linux distributions [25], [26] eschew solving altogether, opting instead to focus on reproducible configurations maintained by humans. Most programmers do not know how to use solvers effectively, and fast, high-quality solver implementations do not exist for new and especially interpreted languages. Moreover, package managers are now fundamental to software ecosystems, and most language communities prefer to write and maintain their core tooling in their own language.

Despite this, developers are starting to realize the need for completeness and well defined dependency resolution semantics [20]. The Python community, plagued by inconsistencies in resolutions done by PIP has recently switched to a new resolver with a proper solver [8]. Dart now uses a custom CDCL SAT solver called PubGrub [24], and Rust’s Cargo [23] package manager is moving towards this approach [27]. However, these solvers use ad hoc techniques baked into the implementations to produce desirable solutions, such as exploring package versions sorted by version number. These are not guaranteed to be optimal, and it is unclear how to add or modify objectives to these types of solvers. In contrast, PACSOLVE makes two new innovations: PACSOLVE allows for

a declarative specification of multiple prioritized optimization objectives, and PACSOLVE changes the problem representation from prior works’ boolean-variable-per-dependency representation based on SAT solving to PACSOLVE’s symbolic graph representation (Section III-C) based on SMT constraints.

Solvers themselves are becoming more accessible through tools like Rosette [5], which makes features of the Z3 [28] SMT solver accessible within regular Racket [29] code, and which we leverage to implement PACSOLVE. Spack [30] makes complex constraints available in a Python DSL, and implements their semantics using Answer Set Programming [31], [32]. APT is moving towards using Z3 to implement more sophisticated dependency semantics [33].

The goal of our work is to further separate concerns away from package manager developers. PACSOLVE focuses on *consistency* criteria and formalizes the *guarantees* that can be offered by package solvers. NPM [21]’s tree-based solver avoids the use of an NP-complete solver by allowing multiple, potentially *inconsistent* versions of the same package in a tree. Tools like Yarn [34], NPM’s audit tool [1], Dependabot [35], Snyk [36] and others [37], [38], [39], [40] attempt to answer various needs of developers by using ad hoc techniques separate from the solving phase, such as deduplication via *hoisting* [41] (Yarn), or post hoc updating of dependencies (NPM’s audit tool). However, these tools run the risk of both correctness bugs and non-optimality in their custom algorithms. PACSOLVE provides the best of all these worlds. It combines the flexibility of multi-version resolution algorithms with the guarantees of complete package solvers and being able to reason about multiple optimization objectives that each speak to a need of developers, while guaranteeing a *minimal* dependency graph.

## VI. DISCUSSION

### A. NPM

By modeling NPM in PACSOLVE and comparing their real-life behavior, we gained valuable insight into NPM’s behavior. As already explored in Section IV, NPM is non-optimal, and it is challenging to see how it could be optimal without implementing a full solver based approach such as MAXNPM. However, NPM does have some lower-hanging fruit that is easier to achieve and would benefit users. First, NPM is not in fact *complete*, in that there are situations where a satisfying solution exists, but NPM fails to find it. Most commonly, a version of a package depends on a dependency which does not exist, and NPM immediately bails out rather than backtracking. This could be implemented with simple backtracking without harming the performance of solves which currently succeed. In addition, Section II-A identifies several shortcomings of the `npm audit fix` tool at the time of our testing. We would suggest incorporating severity of vulnerabilities into the update logic, so that the tool can decide trade-offs between different vulnerabilities. The tool is also unable to downgrade dependencies to remove vulnerabilities, which would be a useful option to have, even if not enabled by default.

---

```

1 (define (cargo-consistent v1 v2)
2   (match `(.v1 .v2)
3     [ `((0 0 ,z1) (0 0 ,z2))          #true]
4     [ `((0 ,y ,z1) (0 ,y ,z2))        (= z1 z2)]
5     [ `((0 ,y1 ,z1) (0 ,y2 ,z2))      #true]
6     [ `(x ,y1 ,z1) (x ,y2 ,z2))      (and (= y1 y2) (= z1 z2))]
7     [_                                 #true]))

```

---

Fig. 8: A consistency function for Cargo.

NPM also contains some subtle behavior regarding *release tags*. NPM version numbers may include release tags, such as `1.2.3-alpha.1`, which MAXNPM fully supports. In particular, a prerelease version can only satisfy a constraint if a sub-term of the constraint with the same semver version also has a prerelease. For example, consider the following constraint (spaces indicate a conjunction):

```
>1.2.3-alpha.3 <1.5.2-alpha.8
```

It may be obvious that versions `1.2.3-alpha.7` and `1.5.2-alpha.6` do satisfy the constraint. However, version `1.3.4` also satisfies the constraint, while version `1.3.4-alpha.7` *does not* satisfy the constraint. A key feature of PACSOLVE which enabled us to easily implement this feature is that the datatype of version numbers ( $\mathcal{V}$ ) is not fixed to be e.g. a 3-tuple of integers, but can be customized, which we leverage to correctly implement NPM’s release tags.

#### B. Future Work: Cargo, Maven and NuGet

Cargo has a notion of *feature flags* which enable conditional compilation. To model this, versions may have the form  $(x, y, z, F)$ , where  $F$  is a set of selected features. To then build a Cargo-compatible dependency solver, we would need to 1) define a constraint satisfaction predicate that ensures that the selected features are a superset of the requested feature, and 2) an objective function that ensures that the minimal number of required features are enabled. In addition, Cargo has a more complex consistency function which allows two versions to be co-installed if they are not semver-compatible. That is easily implemented in PACSOLVE as shown in Figure 8.

Maven and NuGet resolve all dependencies to exactly one version for each package. However, when conflicts arise, instead of backtracking to an older version, they *ignore all but the closest package constraint to the root*. This is arguably unsound, but it is possible to model this behavior in PACSOLVE by treating constraints as soft constraints that are weighted by their distance from the root.

#### C. Threats to Validity

a) *External Validity*: The projects that we used in our evaluation may not be representative of the entire ecosystem of NPM packages. We select the 1,000 most popular projects, and report performance as a distribution over this entire dataset, including a discussion of outliers. Given the number of projects that we used in our evaluation and their popularity, we believe that MAXNPM is quite likely to be helpful for improving package management in real-world

scenarios. We describe PACSOLVE as a unifying framework for implementing dependency solvers, however we only use PACSOLVE to implement MAXNPM. Future work should empirically validate PACSOLVE’s efficacy in other ecosystems.

b) *Internal Validity*: MAXNPM, PACSOLVE, and the tools that we build upon may have bugs that impact our results. To verify that differences between NPM and MAXNPM are not due to bugs in MAXNPM, we carefully analyzed the cases where MAXNPM and NPM diverged in their solution, and we walkthrough some example cases in Section IV. Additionally, we have carefully written a suite of unit tests for PACSOLVE.

c) *Construct Validity*: We evaluate MAXNPM’s relative performance to NPM when optimizing for several different objective criteria. However, it is possible that these criteria are not meaningful to developers. For example, when comparing MAXNPM and `npm audit fix` in reducing vulnerabilities, we use the aggregate vulnerability scores (CVSS) to rank the tools. However, in practice, these scores may not directly capture the true severity of a vulnerable dependency in the context of a particular application. MAXNPM does however allow for potential customization of constraints to fit the developer’s needs. Future work should involve user studies, observing the direct impact of PACSOLVE-based solvers (including MAXNPM, and implementations for other ecosystems) on developers.

## VII. CONCLUSION

We present PACSOLVE, a semantics of dependency solving that we use to highlight the essential features and variation within the package manager design space. We use PACSOLVE to implement MAXNPM, a drop-in replacement for NPM that allows the user to customize dependency solving with a variety of global objectives and consistency criteria. Using MAXNPM, developers can optimize dependency resolution to achieve goals that NPM is unable to, such as: reduce the presence of vulnerabilities, resolve newer packages and reduce bloat. We evaluate MAXNPM on the top 1,000 packages in the NPM ecosystem, finding that our prototype introduces a median overhead of less than two seconds. We find that MAXNPM produces solutions with fewer dependencies and newer dependencies for many packages. For future work, we hope to use PACSOLVE to build new dependency solvers for other package managers as well.

## VIII. DATA AVAILABILITY

Our artifact is available under a CC-BY-4.0 license [42] and consists of 1) the implementations of PACSOLVE and MAXNPM, 2) the *Top1000* and *Vuln715* datasets, and 3) scripts to reproduce our results. All of our code is also available on GitHub [43], and MAXNPM can be easily installed with `npm install -g maxnpm`.

## ACKNOWLEDGMENTS

We thank Northeastern Research Computing, especially Greg Shomo, for computing resources and technical support.

## REFERENCES

- [1] NPM, “npm-audit,” <https://docs.npmjs.com/cli/v9/commands/npm-audit>, 2023.
- [2] R. Roemer, “Finding and fixing duplicates in webpack with Inspectpack,” [formidable.com/blog/2018/finding-webpack-duplicates-with-inspectpack-plugin/](https://formidable.com/blog/2018/finding-webpack-duplicates-with-inspectpack-plugin/). Accessed Aug 29 2022, 2022.
- [3] S. Neumann, “Similar files are collapsed (breaking react-bootstrap),” [github.com/parcel-bundler/parcel/issues/3523](https://github.com/parcel-bundler/parcel/issues/3523). Accessed Aug 29 2022, 2019.
- [4] W. Contributors, “Mark the file as side-effect-free,” [webpack.js.org/guides/tree-shaking/#mark-the-file-as-side-effect-free](https://webpack.js.org/guides/tree-shaking/#mark-the-file-as-side-effect-free). Accessed Aug 29 2022, 2022.
- [5] E. Torlak and R. Bodik, “Growing solver-aided languages with rosette,” in *Proceedings of the 2013 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming & Software*, ser. Onward! 2013. New York, NY, USA: Association for Computing Machinery, 2013, p. 135–152. [Online]. Available: <https://doi.org/10.1145/2509578.2509586>
- [6] C. Tucker, D. Shuffelton, R. Jhala, and S. Lerner, “Opium: Optimal package install/uninstall manager,” in *Proceedings of the 29th International Conference on Software Engineering*, ser. ICSE ’07. USA: IEEE Computer Society, 2007, pp. 178–188.
- [7] R. Di Cosmo, “EDOS deliverable WP2-D2.1: Report on Formal Management of Software Dependencies,” INRIA, Tech. Rep., May 15 2005, hal-00697463.
- [8] Python Software Foundation, “New pip resolver to roll out this year,” Online, March 23 2020, <https://pyfound.blogspot.com/2020/03/new-pip-resolver-to-roll-out-this-year.html>.
- [9] M. Broy and M. Wirsing, “On the algebraic specification of nondeterministic programming languages,” in *Colloquium on Trees in Algebra and Programming*. Springer, 1981, pp. 162–179.
- [10] C. Contributors, “Conda performance,” [docs.conda.io/projects/conda/en/latest/user-guide/concepts/conda-performance.html](https://docs.conda.io/projects/conda/en/latest/user-guide/concepts/conda-performance.html). Accessed Sep 1 2022, 2022.
- [11] A. Van Der Hoek, R. S. Hall, D. Heimbigner, and A. L. Wolf, “Software release management,” *ACM SIGSOFT Software Engineering Notes*, vol. 22, no. 6, pp. 159–175, 1997.
- [12] M. Ewing and E. Troan, “RPM Timeline,” Online, 1995, <https://rpm.org/timeline.html>.
- [13] J. Gunthorpe, “APT User’s Guide,” Online, 1998, <https://www.debian.org/doc/manuals/apt-guide/>.
- [14] F. Mancinelli, J. Boender, R. di Cosmo, J. Vouillon, B. Durak, X. Leroy, and R. Treinen, “Managing the complexity of large free and open source package-based software distributions,” in *21st IEEE/ACM International Conference on Automated Software Engineering (ASE’06)*, 2006, pp. 199–208.
- [15] P. Abate, R. Di Cosmo, R. Treinen, and S. Zacchiroli, “Dependency solving: a separate concern in component evolution management,” *Journal of Systems and Software*, vol. 85, no. 10, pp. 2228–2240, 2012.
- [16] P. Abate, R. Di Cosmo, R. Treinen, and S. Zacchiroli, “A modular package manager architecture,” *Information and Software Technology*, vol. 55, no. 2, pp. 459 – 474, 2013, special Section: Component-Based Software Engineering (CBSE), 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950584912001851>
- [17] C. Michel and M. Rueher, “Handling software upgradeability problems with MILP solvers,” in *Proceedings First International Workshop on Logics for Component Configuration, LoCoCo 2010, Edinburgh, UK, 10th July 2010*, ser. EPTCS, I. Lynce and R. Treinen, Eds., vol. 29, 2010, pp. 1–10. [Online]. Available: <https://doi.org/10.4204/EPTCS.29.1>
- [18] J. Argelich, D. Le Berre, I. Lynce, J. P. M. Silva, and P. Rapiçault, “Solving linux upgradeability problems using boolean optimization,” in *Proceedings First International Workshop on Logics for Component Configuration, LoCoCo 2010, Edinburgh, UK, 10th July 2010*, ser. EPTCS, I. Lynce and R. Treinen, Eds., vol. 29, 2010, pp. 11–22. [Online]. Available: <https://doi.org/10.4204/EPTCS.29.2>
- [19] M. Gebser, R. Kaminski, and T. Schaub, “aspcud: A linux package configuration tool based on answer set programming,” *Electronic Proceedings in Theoretical Computer Science*, vol. 65, pp. 12–25, Aug 2011.
- [20] P. Abate, R. Di Cosmo, G. Gousios, and S. Zacchiroli, “Dependency solving is still hard, but we are getting better at it,” in *2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER)*. IEEE, 2020, pp. 547–551.
- [21] I. Z. Schlueter, “NPM,” Online, September 2009, <https://github.com/npm/npm>.
- [22] I. Bicking, “pip: Package Install tool for Python,” April 2011, [github.com/pypa/pip](https://github.com/pypa/pip).
- [23] “Cargo: The Rust package manager,” Online, March 2014, <https://github.com/rust-lang/cargo>.
- [24] N. Weizenbaum, “PubGrub: Next-Generation Version Solving,” <https://medium.com/@nex3/pubgrub-2fb6470504f>, April 2 2018.
- [25] E. Dolstra, M. de Jonge, and E. Visser, “Nix: A Safe and Policy-Free System for Software Deployment,” in *Proceedings of the 18th Large Installation System Administration Conference (LISA XVIII)*, ser. LISA ’04. Berkeley, CA, USA: USENIX Association, 2004, pp. 79–92.
- [26] L. Courtès and R. Wurmus, “Reproducible and User-Controlled Software Environments in HPC with Guix,” in *2nd International Workshop on Reproducibility in Parallel Computing (RepPar)*, Vienne, Austria, Aug. 2015. [Online]. Available: <https://hal.inria.fr/hal-01161771>
- [27] “PubGrub version solving algorithm implemented in Rust,” Online, 2020, <https://github.com/pubgrub-rs/pubgrub>.
- [28] L. De Moura and N. Bjørner, “Z3: An efficient smt solver,” in *International conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 2008, pp. 337–340.
- [29] M. Felleisen, R. B. Fidler, M. Flatt, S. Krishnamurthi, E. Barzilay, J. McCarthy, and S. Tobin-Hochstadt, “The racket manifesto,” in *1st Summit on Advances in Programming Languages (SNAPL 2015)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2015.
- [30] T. Gamblin, M. P. LeGendre, M. R. Collette, G. L. Lee, A. Moody, B. R. de Supinski, and W. S. Futral, “The Spack Package Manager: Bringing order to HPC software chaos,” in *Supercomputing 2015 (SC’15)*, Austin, Texas, November 15-20 2015, ILNL-CONF-669890.
- [31] M. Gebser, B. Kaufmann, R. Kaminski, M. Ostrowski, T. Schaub, and M. Schneider, “Potassco: The potsdam answer set solving collection,” *AI Communications*, vol. 24, no. 2, pp. 107–124, 2011.
- [32] T. Gamblin, “Spack’s new Concretizer: Dependency solving is more than just SAT!” in *Free and Open source Software Developers’ European Meeting (FOSDEM’20)*, Brussels, Belgium, February 1 2020.
- [33] J. A. Klode, “APT Z3 Solver Basics,” Online, November 21 2021, <https://blog.jak-linux.org/2021/11/21/apt-z3-solver-basics/>.
- [34] “Yarn: Yet Another Resource Negotiator (javascript package manager),” 2022, <https://github.com/yarnpkg/yarn>.
- [35] GitHub, “Dependabot,” <https://github.com/features/security/>, 2023.
- [36] Snyk, “Snyk,” <https://snyk.io>, 2023.
- [37] C. Soto-Valero, N. Harrand, M. Monperrus, and B. Baudry, “A comprehensive study of bloated dependencies in the maven ecosystem,” *Empirical Software Engineering*, vol. 26, 2021. [Online]. Available: <https://link.springer.com/article/10.1007/s10664-020-09914-8#citeas>
- [38] S. E. Ponta, W. Fischer, H. Plate, and A. Sabetta, “The used, the bloated, and the vulnerable: Reducing the attack surface of an industrial application,” in *2021 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, 2021, pp. 555–558.
- [39] C. Soto-Valero, T. Durieux, and B. Baudry, “A longitudinal analysis of bloated java dependencies,” in *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ser. ESEC/FSE 2021. New York, NY, USA: Association for Computing Machinery, 2021, p. 1021–1031. [Online]. Available: <https://doi.org/10.1145/3468264.3468589>
- [40] J. Latendresse, S. Mujahid, D. E. Costa, and E. Shihab, “Not all dependencies are equal: An empirical study on production dependencies in npm,” in *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*, ser. ASE ’22. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3551349.3556896>
- [41] T. Maier, “A guide to understanding how Yarn hoists dependencies and handles conflicting packages,” Online, November 27 2021, <https://maier.tech/posts/a-guide-to-understanding-how-yarn-hoists-dependencies-and-handles-conflicting-packages>.
- [42] D. Pinckney, F. Cassano, A. Guha, J. Bell, M. Culpo, and T. Gamblin, “Artifact for Flexible and Optimal Dependency Management via Max-SMT,” <https://doi.org/10.5281/zenodo.7554407>, 2023.
- [43] D. Pinckney and F. Cassano, “PacSolve,” <https://github.com/donald-pinckney/pacsolve>, 2023.