LWeb: Information Flow Security for Multi-tier Web Applications

JAMES PARKER, University of Maryland, USA
NIKI VAZOU*, IMDEA Software Institute, Spain
MICHAEL HICKS, University of Maryland, USA

This paper presents LWeb, a framework for enforcing label-based, information flow policies in database-using web applications. In a nutshell, LWeb marries the LIO Haskell IFC enforcement library with the Yesod web programming framework. The implementation has two parts. First, we extract the core of LIO into a monad transformer (LMonad) and then apply it to Yesod's core monad. Second, we extend Yesod's table definition DSL and query functionality to permit defining and enforcing label-based policies on tables and enforcing them during query processing. LWeb's policy language is expressive, permitting dynamic per-table and per-row policies. We formalize the essence of LWeb in the λ_{LWeb} calculus and mechanize the proof of noninterference in Liquid Haskell. This mechanization constitutes the first metatheoretic proof carried out in Liquid Haskell. We also used LWeb to build a substantial web site hosting the *Build it, Break it, Fix it* security-oriented programming contest. The site involves 40 data tables and sophisticated policies. Compared to manually checking security policies, LWeb imposes a modest runtime overhead of between 2% to 21%. It reduces the trusted code base from the whole application to just 1% of the application code, and 21% of the code overall (when counting LWeb too).

CCS Concepts: • **Software and its engineering** → **Semantics**; • **Security and privacy** → *Logic and verification*; *Web application security*;

Additional Key Words and Phrases: security, information flow control, metatheory, Liquid Haskell, Haskell

ACM Reference Format:
James Parker, Niki Vazou, and Michael Hicks. 2019. LWeb: Information Flow Security for Multi-tier Web Applications. *Proc. ACM Program. Lang.* 3, POPL, Article 75 (January 2019), 30 pages. https://doi.org/10.1145/3290388

1 INTRODUCTION

Modern web applications must protect the confidentiality and integrity of their data. Employing access control and/or manual, ad hoc enforcement mechanisms may fail to block illicit information flows between components, e.g., from database to server to client. Information flow control (IFC) [Sabelfeld and Myers 2003] policies can govern such flows, but enforcing them poses practical problems. Static enforcement (e.g., by typing [Chong et al. 2007a,b; Myers 1999; Pottier and Simonet 2003; Schoepe et al. 2014] or static analysis [Arzt et al. 2014; Hammer and Snelting 2009; Johnson et al. 2015]) can produce too many false alarms, which hamper adoption [King et al. 2008]. Dynamic enforcement [Austin and Flanagan 2012; Chudnov and Naumann 2015; Roy et al. 2009; Tromer and Schuster 2016; Yang et al. 2016] is more precise but can impose high overheads.

*Vazou was at the University of Maryland while this work was carried out.

Authors' addresses: James Parker, Department of Computer Science, University of Maryland, USA; Niki Vazou, IMDEA Software Institute, Spain; Michael Hicks, Department of Computer Science, University of Maryland, USA.

2475-1421/2019/1-ART75
https://doi.org/10.1145/3290388

A promising solution to these problems is embodied in the LIO system [Stefan et al. 2011] for Haskell. LIO is a drop-in replacement for the Haskell IO monad, extending IO with an internal *current label* and *clearance label*. Such labels are lattice ordered (as is typical [Denning 1976]), with the degenerate case being a secret (high) label and public (low) one. LIO's current label constitutes the least upper bound of the security labels of all values read during the current computation. Effectful operations such as reading/writing from stable storage, or communicating with other processes, are checked against the current label. If the operation's security label (e.g., that on a channel being written to) is lower than the current label, then the operation is rejected as potentially insecure. The clearance serves as an upper bound that the current label may never cross, even prior to performing any I/O, so as to reduce the chance of side channels. Haskell's clear, type-enforced separation of pure computation from effects makes LIO easy to implement soundly and efficiently, compared to other dynamic enforcement mechanisms.

This paper presents `LWeb`, an extension to `LIO` that aims to bring its benefits to Haskell-based web applications. This paper presents the three main contributions of our work.

First, we present an extension to a core LIO formalism with support for database transactions. Each table has a label that protects its length. In our implementation we use DC labels [Stefan et al. 2012], which have both confidentiality and integrity components. The confidentiality component of the table label controls who can query it (as the result may reveal something about the table's length), and the integrity component controls who can add or delete rows (since both may change the length). In addition, each row may have a more refined policy to protect its contents. The label for a field in a row may be specified as a function of other fields in the same row (those fields are protected by a specific, global label). This allows, for example, having a row specifying a user and some sensitive user data; the former can act as a label to protect the latter.

We mechanized our formalism in Liquid Haskell [Vazou et al. 2014] and proved that it enjoys noninterference. Our development proceeds in two steps: a core `LIO` formalism called λ_{LIO} (§ 3), and an extension to it, called λ_{LWeb}, that adds database operations (§ 4). The mechanization process was fruitful: it revealed two bugs in our original rules that constituted real leaks. Moreover, this mechanization constitutes the largest-ever development in Liquid Haskell and is the first Liquid Haskell application to prove a language metatheory (§ 5).

As our next contribution, we describe a full implementation of `LWeb` in Haskell as an extension to the `Yesod` web programming framework (§ 2 and § 6). Our implementation was carried out in two steps. First, we extracted the core label tracking functionality of `LIO` into a monad transformer called `LMonad` so that it can be layered on monads other than `IO`. For `LWeb`, we layered it on top of the `Handler` monad provided by the `Yesod`. This monad encapsulates mechanisms for client/server HTTP communications and database transactions, so layering `LMonad` on top of `Handler` provides the basic functionality to enforce security. Then we extended `Yesod`'s database API to permit defining label-based information flow policies, generalizing the approach from our formalism whereby each row may have many fields, each of which may be protected by other fields in the same row. We support simple key/value lookups and more general SQL queries, extending the Esqueleto framework [Esqueleto 2018]. We use Template Haskell [Sheard and Jones 2002] to insert checks that properly enforce policies in our extension.

Finally, we describe our experience using `LWeb` to build a substantial web site hosting the Build it, Break it, Fix it (BIBIFI) security-oriented programming contest [Ruef et al. 2016] hosted at https://builditbreakit.org (§ 7). This site has been used over the last few years to host more than a dozen contests involving hundreds of teams. It consists of 11500+ lines of Haskell and manages data stored in 40 database tables. The site has a variety of roles (participants, teams, judges, admins) and policies that govern their various privileges. When we first deployed this contest, it lacked `LWeb` support, and we found it had authorization bugs. Retrofitting it with `LWeb` was straightforward and

eliminated those problems, reducing the trusted computing base from the entire application to just 80 lines of its code (1%) plus the LWeb codebase (for a total of 21%). LWeb imposes modest overhead on BIBIFI query latencies—experiments show between 2% and 21% (§ 8).

LWeb is not the first framework to use IFC to enforce database security in web applications. Examples of prior efforts include SIF/Swift [Chong et al. 2007a,b], Jacqueline [Yang et al. 2016], Hails [Giffin et al. 2017, 2012], SELinks [Corcoran et al. 2009], SeLINQ [Schoepe et al. 2014], UrFlow [Chlipala 2010], and IFDB [Schultz and Liskov 2013]. LWeb distinguishes itself by providing end-to-end IFC security (between/across server and database), backed by a formal proof (mechanized in Liquid Haskell), for a mature, full-featured web framework (Yesod) while supporting expressive policies (e.g., where one field can serve as the label of another) and efficient queries (a large subset of SQL). The IFC checks needed during query processing were tricky to get right—our formalization effort uncovered bugs in our original implementation by which information could leak owing to the checks themselves. § 9 discusses related work in detail.

The code for LWeb and its mechanized proof are freely available.

2 OVERVIEW

The architecture of LWeb is shown in fig. 1. Database queries/updates precipitated by user interactions are processed by the LMonad component, which constitutes the core of LIO and confirms that label-based security policies are not violated. Then, the queries/updates are handled via Yesod, where the results continue to be subject to policy enforcement by LMonad.

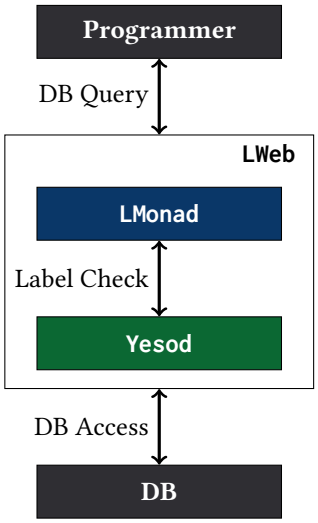


Fig. 1. Structure of LWeb.

2.1 Label-Based Information Flow Control with LIO

We start by presenting LIO [Stefan et al. 2011] and how it is used to enforce noninterference for label-based information flow policies.

Labels and noninterference. As a trivial security label, consider a datatype with constructors Secret and Public. Protected data is assigned a label, and an IFC system ensures that Secret-labeled data can only be learned by those with Secret-label privilege or greater. The label system can be generalized to any lattice [Denning 1976] where IFC is checked using the lattice's partial order relation $\sqsubseteq$. Such a system enjoys *noninterference* [Goguen and Meseguer 1982] if an adversary with privileges at label l1 can learn nothing about data labeled with l2 where $l2 \not\sqsubseteq l1$.

In fig. 2 we define the label interface as the type class Label that defines the bottom (least protected) label, least upper bound (join, $\sqcup$) of two labels, the greatest lower bound (meet, $\sqcap$), and whether one label can flow to ($\sqsubseteq$) another, defining a partial ordering. Instantiating this type class for Public and Secret would set Public as the bottom label and Public $\sqsubset$ Secret (with join and meet operations to match).

```
class Eq a ⇒ Label a
  where
  ⊥    :: a
  (⊔) :: a → a → a
  (⊓) :: a → a → a
  (⊑) :: a → a → Bool
```

Fig. 2. The Label class

The LIO monad. LIO enforces IFC on labeled data using dynamic checks. The type LIO l a denotes a monadic computation that returns a value of type a at label l. LIO provides two methods to label and unlabel data.

```
label   :: (Label l) ⇒ l → a → LIO l (Labeled l a)
```

```
Friends <⊥,Const Admin>
 user1 Text <⊥,Const Admin>
 user2 Text <⊥,Const Admin>
 date  Text <Field User1 ⊓ Field User2,Const Admin>
```

Fig. 3. Example LWeb database table definition. The green is Yesod syntax and the blue is the *LWeb* policy.

```
unlabel :: (Label l) ⇒ Labeled l a → LIO l a
```

The method `label l v` takes as input a label and some data and returns a `Labeled` value, i.e., the data `v` marked with the label `l`. The method `unlabel v` takes as input a labeled value and returns just its data. The `LIO` monad maintains an ambient label—the *current label* `lc`—that represents the label of the current computation. As such, labelling and unlabelling a value affects `lc`. In particular, `unlabel v` updates `lc` by joining it to `v`'s label, while `label l v` is only permitted if `lc` $\sqsubseteq$ `l`, i.e., the current label can flow to `l`. If this check fails, `LIO` raises an exception.

As an example, on the left, a computation with current label `Public` labels data `"a secret"` as `Secret`, preserving the same current label, and then unlabels the data, thus raising the current label to `Secret`. On the right, a computation with current label `Secret` attempts to label data as `Public`, which fails, since the computation is already tainted with (i.e., dependent on) secret data.

```
-- lc := Public
v ← label Secret "a secret"
-- ok: Public ⊑ Secret and lc := Public
x ← unlabel v
-- lc := Secret
```

```
-- lc := Secret
v ← label Public "public"
-- exception: Secret ⋢ Public
```

`LIO` also supports labeled mutable references, and a scoping mechanism for temporarily (but safely) raising the current label until a computation completes, and then restoring it. `LIO` also has what is called the *clearance* label that serves as an upper bound for the current label, and thus can serve to identify potentially unsafe computations sooner.

A normal Haskell program can run an `LIO` computation via `runLIO`, whose type is as follows.

```
runLIO :: (Label l) ⇒ LIO l a → IO a
```

Evaluating `runLIO m` initializes the current label to $\bot$ and computes `m`. The returned result is an `IO` computation, since `LIO` allows `IO` interactions, e.g., with a file system. If any security checks fail, `runLIO` throws an exception.

2.2 Yesod

`Yesod` [Snoyman 2018] is mature framework for developing type-safe and high performance web applications in Haskell. In a nutshell, `LWeb` adds `LIO`-style support to `Yesod`-based web applications, with a focus on supporting database security policies.

The green part of fig. 3 uses `Yesod`'s domain specific language (DSL) to define the table `Friends`. The table has three `Text`[1] fields corresponding to two users (`user1` and `user2`) and the date of their friendship. A primary key field with type `FriendsId` is also automatically added. In § 2.3 we explain how the blue part of the definition is used for policy enforcement.

`Yesod` uses Template Haskell [Sheard and Jones 2002] to generate, at compile time, a database schema from such table definitions. These are the Haskell types that `Yesod` generates for the `Friends` table.

[1] `Text` is an efficient Haskell string type.

```
data FriendsId = FriendsId Int
data Friends = Friends { friendsUser1 :: Text, friendsUser2 :: Text
                       , friendsDate  :: Text }
```

Note that though each row has a key of type `FriendsId`, it is elided from the `Friends` data record. Each generated key type is a member of the `Key` type family; in this case `Key Friends` is a type alias for `FriendsId`.

`Yesod` provides an API to define and run queries. Here is a simplified version of this API.

```
runDB  :: YesodDB a → Handler a
get    :: Key v → YesodDB (Maybe v)
insert :: v     → YesodDB (Key v)
delete :: Key v → YesodDB ()
update :: Key v → [Update v] → YesodDB ()
```

The type alias `YesodDB a` denotes the monadic type of a computation that queries (or updates) the database. The function `runDB` runs the query argument on the database. `Handler` is `Yesod`'s underlying monad used to respond to HTTP requests. The functions `get`, `insert`, `delete`, and `update` generate query computations. For example, we can query the database for the date of a specific friendship using `get`.

```
getFriendshipDate :: FriendsId → Handler (Maybe Text)
getFriendshipDate friendId = do
    r ← runDB (get friendId)
    return (friendsDate <$> r)
```

`Yesod` also supports more sophisticated SQL-style queries via an interface called Esqueleto [Esqueleto 2018]. Such queries may include inner and outer joins, conditionals, and filtering.

2.3 LWeb: Yesod with LIO

`LWeb` extends `Yesod` with `LIO`-style IFC enforcement. The implementation has two parts. As a first step, we generalize `LIO` to support an arbitrary underlying monad by making it a *monad transformer*, applying it to `Yesod`'s core monad. Then we extend `Yesod` operations to incorporate label-based policies that work with this extended monad.

`LMonad`*: LIO as a monad transformer.* `LMonad` generalizes the underlying `IO` monad of `LIO` to *any* monad `m`. In particular, `LMonad` is a monad transformer `LMonadT l m` that adds the IFC operations to the underlying monad `m`, rather than making it specific to the `IO` monad.

```
label     :: (Label l, Monad m) ⇒ l → a → LMonadT l m (Labeled l a)
unlabel   :: (Label l, Monad m) ⇒ Labeled l a → LMonadT l m a
runLMonad :: (Label l, Monad m) ⇒ LMonadT l m a → m a
```

`LMonadT` is implemented as a state monad transformer that tracks the current label. Computations that run in the underlying `m` monad cannot be executed directly due to Haskell's type system. Instead, safe variants that enforce IFC must be written so that they can be executed in `LMonadT l m`. Thus, the `LIO` monad is an instantiation of the monad variable `m` with `IO`: `LIO l = LMonadT l IO`. For `LWeb` we instantiate `LMonadT` with `Yesod`'s `Handler` monad.

```
type LHandler l a = LMonadT l Handler a
```

Doing this adds information flow checking to `Yesod` applications, but it still remains to define policies to be checked. Thus we extend `Yesod` to permit defining label-based policies on database schemas, and to enforce those policies during query processing.

Label-annotated database schemas. LWeb labels are based on DC labels [Stefan et al. 2012], which have the form <l , r>, where the left protects the *confidentiality* and the right protects the *integrity* of the labeled value. Integrity lattices are dual to confidentiality lattices. They track who can influence the construction of a value.

Database policies are written as label annotations p on table definitions, following this grammar:

```
p := <l, l>
l := Const c | Field f | Id | ⊤ | ⊥ | l ⊓ l | l ⊔ l
```

Here, c is the name of a data constructor and f is a field name. A database policy consists of a single *table label* and one label for each field in the database. We explain these by example.

The security labels of the Friends table are given by the blue part of fig. 3. The first line's label `Friends <⊥,Const Admin>` defines the table label, which protects the *length* of the table. This example states that anyone can learn the length of the table (e.g., by querying it), but only the administrator can change the length (i.e., by adding or removing entries). LWeb requires the table label to be constant, i.e., it may not depend on run-time entries of the table. Allowing it to do so would significantly complicate enforcing noninterference.

The last line `date Text <Field User1 ⊓ Field User2,Const Admin>` defines that either of the users listed in the first two fields can read the date field but only the administrator can write it. This label is *dynamic*, since the values of the user1 and user2 fields may differ from row to row. We call fields, like user1 and user2, which are referenced in another field's label annotation, *dependency fields*. When a field's label is not given explicitly, the label <⊥ , ⊤> is assumed. To simplify security enforcement, LWeb requires the label of a dependency field to be constant and flow into (be bounded by) the table label. For user1 and user2 this holds since their labels match the table's label.

The invariants about the table label and the dependency field labels are enforced by a compile-time check, when processing the table's policy annotations. Note that Labeled values may not be directly stored in the database as there is no way to directly express such a type in a source program. Per fig. 3, field types like Text, Bool, and Int are allowed, and their effective label is indicated by annotation, rather than directly expressed in the type.[2]

Policy enforcement. LWeb enforces the table-declared policies by providing wrappers around each Yesod database API function.

```
runDB  :: Label l ⇒ LWebDB l a → LHandler l a
get    :: Label l ⇒ Key v → LWebDB l (Maybe v)
insert :: Label l ⇒ v     → LWebDB l (Key v)
delete :: Label l ⇒ Key v → LWebDB l ()
update :: Label l ⇒ Key v → [Update v] → LWebDB l ()
```

Now the queries are modified to return LWebDB computations that are evaluated (using runDB) inside the LHandler monad. For each query operation, LWeb wraps the underlying database query with information flow control checks that enforce the defined policies. For instance, if x has type FriendsId, then `r ← runDB $ get x` joins the current label with the label of the selected row, here `user1 ⊓ user2`.

LWeb also extends IFC checking to advanced SQL queries expressed in Esqueleto [Esqueleto 2018]. As explained in § 6, LWeb uses a DSL syntax, as a lsql quasiquotation, to wrap these queries with IFC checks. For example, the following query joins the Friends table with a User table:

```
rs ← runDB [lsql|select ★ from Friends inner join User on Friends.user1 == User.id|]
```

[2]The formalism encodes all of these invariants with refinement types in the database definition.

```
class Label l where
  (⊑) :: l → l → Bool
  (⊓) :: l → l → l
  (⊔) :: l → l → l
  ⊥   :: l

  lawBot             :: l:l  → { ⊥ ⊑ l }
  lawFlowReflexivity  :: l:l  → { l ⊑ l }
  lawFlowAntisymmetry :: l1:l → l2:l → { (l1 ⊑ l2 ∧ l2 ⊑ l1) ⇒ l1 == l2 }
  lawFlowTransitivity :: l1:l → l2:l → l3:l → { (l1 ⊑ l2 ∧ l2 ⊑ l3) ⇒ l1 ⊑ l3 }

  lawMeet :: z:l → l1:l → l2:l → l:l
          → { z == l1 ⊓ l2 ⇒ z ⊑ l1 ∧ z ⊑ l2 ∧ (l ⊑ l1 ∧ l ⊑ l2 ⇒ l ⊑ z) }
  lawJoin :: z:l → l1:l → l2:l → l:l
          → { z == l1 ⊔ l2 ⇒ l1 ⊑ z ∧ l2 ⊑ z ∧ (l1 ⊑ l ∧ l2 ⊑ l ⇒ z ⊑ l) }
```

Fig. 4. `Label` type class extended with `law*` methods to define the lattice laws as refinement types.

3 MECHANIZING NONINTERFERENCE OF LIO IN LIQUID HASKELL

A contribution of this work is a formalization of LWeb's extension to LIO to support database security policies, along with a proof that this extension satisfies (termination insensitive) noninterference. We mechanize our formalization in Liquid Haskell [Vazou et al. 2014], an SMT-based refinement type checker for Haskell programs. Liquid Haskell permits refinement type specifications on Haskell source code. It converts the code into SMT queries to validate that the code satisfies the specifications. Our mechanized formalization and proof of noninterference constitutes the first significant metatheoretical mechanization carried out in Liquid Haskell.

We present our mechanized LWeb formalism in two parts. In this section, we present λ_{LIO}, a formalization and proof of noninterference for LIO. The next section presents λ_{LWeb}, an extension of λ_{LIO} that supports database operations. Our Liquid Haskell mechanization defines λ_{LIO}'s syntax and operational semantics as Haskell definitions, as a definitional interpreter. We present them the same way in this paper, rather than reformatting them as mathematical inference rules. Metatheoretic properties are expressed as refinement types, following Vazou et al. [2017, 2018], and proofs are Haskell functions with these types (checked by the SMT solver). We assess our experience using Liquid Haskell for metatheory in comparison to related approaches in § 5.

3.1 Security Lattice as a Type Class

Figure 4 duplicates the `Label` class definition of Figure 2 but extends it with several methods that use refinement types to express properties of lattices that labels are expected to have.

Partial order. The method `(⊑)` defines a partial order for each `Label` element. That is, `(⊑)` is reflexive, antisymmetric, and transitive, as respectively encoded by the refinement types of the methods `lawFlowReflexivity`, `lawFlowAntisymmetry`, and `lawFlowTransitivity`. For instance, `lawFlowReflexivity` is a method that takes a label `l` to a Haskell unit (i.e., `l → ()`). This type is refined to encode the reflexivity property `l:l → {v:() | l ⊑ l }` and further simplifies to ignore the irrelevant `v:()` part as `l:l → { l ⊑ l }`. With that refinement, application of `lawFlowReflexivity` to a concrete label `l` gives back a proof that `l` can flow to itself (i.e., `l ⊑ l`). At an instance definition of the class `Label`, the reflexivity proof needs to be explicitly provided.

```
data Program l = Pg { pLabel :: l, pTerm :: Term l } | PgHole

data Term l
 -- pure terms
 = TUnit | TInt Int | TLabel l | TLabeled l (Term l) | TLabelOf (Term l)
 | TVar Var | TLam Var (Term l) | TApp (Term l) (Term l) | THole | ...
 -- monadic terms
 | TBind   (Term l) (Term l) | TReturn  (Term l) | TGetLabel | TLIO (Term l)
 | TTLabel (Term l) (Term l) | TUnlabel (Term l) | TException
 | TToLabeled (Term l) (Term l)
```

Fig. 5. Syntax of λ_{LIO}.

Lattice. Similarly, we refine the `lawMeet` method to define the properties of the `(⊓)` lattice operator. Namely, for all labels `l1` and `l2`, we define `z == l1 ⊓ l2` so that (i) `z` can flow to `l1` and `l2` (`z ⊑ l1 ∧ z ⊑ l2`) and (ii) all labels that can flow to `l1` and `l2`, can also flow to `z` (`forall l. l ⊑ l1 ∧ l ⊑ l2 ⇒ l ⊑ z`). Dually, we refine the `lawJoin` method to describe `l1 ⊔ l2` as the minimum label that is greater than `l1` and `l2`.

Using the lattice laws. The lattice laws are class methods, which can be used for each `l` that satisfies the `Label` class constraints. For example, we prove that for all labels `l1`, `l2`, and `l3`, `l1 ⊔ l2` cannot flow into `l3` *iff* `l1` and `l2` cannot both flow into `l3`.

```
joinIff :: Label l ⇒ l1:l → l2:l → l3:l → {l1 ⊑ l3 ∧ l2 ⊑ l3 ⇔ (l1 ⊔ l2) ⊑ l3}
joinIff l1 l2 l3 = lawJoin (l1 ⊔ l2) l1 l2 l3 ? lawFlowTransitivity l1 l2 l3
```

The theorem is expressed as a Haskell function that is given three labels and returns a unit value refined with the desired property. The proof proceeds by calling the laws of join and transitivity, combined with the proof combinator `(?)` that ignores its second argument (i.e., defined as `x ? _ = x`) while passing the refinements of both arguments to the SMT solver. The contrapositive step is automatically enforced by refinement type checking, using the SMT solver.

3.2 λ_{LIO}: Syntax and Semantics

Now we present the syntax and operational semantics of λ_{LIO}.

3.2.1 Syntax. Figure 5 defines a program as either an actual program (`Pg`) with a current label `pLabel` under which the program's term `pTerm` is evaluated, or as a hole (`PgHole`). The hole is not a proper program; it is used for to define adversary observability when proving noninterference (§ 3.3). We omit the clearance label in the formalism as a simplification since its rules are straightforward (when the current label changes, check that it flows into the clearance label). Terms are divided into *pure* terms whose evaluation is independent of the current label and *monadic* terms, which either manipulate or whose evaluation depends on the current label.

Pure terms. Pure terms include unit `TUnit`, integers `TInt i` for some Haskell integer `i`, and the label value `TLabel l`, where `l` is some instance of the labeled class of Figure 4. The labeled value `TLabeled l t` wraps the term `t` with the label `l`. The term `TLabelOf t` returns the label of the term `t`, if `t` is a labeled term. Pure terms include the standard lambda calculus terms for variables (`TVar`), application (`TApp`) and abstraction (`TLam`). Finally, similar to programs, a hole term (`THole`) is required for the meta-theory. It is straightforward to extend pure terms to more interesting calculi.

```
eval :: Label l ⇒ Program l → Program l
eval (Pg lc (TBind t1 t2))
 | Pg lc' (TLIO t1') ← eval* (Pg lc t1)
 = Pg lc' (TApp t2 t1')

eval (Pg lc (TReturn t))
  = Pg lc (TLIO t)

eval (Pg lc TGetLabel)
  = Pg lc (TReturn (TLabel lc))

eval (Pg lc (TTLabel (TLabel l) t))
  | lc ⊑ l
  = Pg lc (TReturn (TLabeled l t))
  | otherwise
  = Pg lc TException

evalTerm :: Label l ⇒ Term l → Term l
evalTerm (TLabelOf (TLabeled l _))
  = TLabel l
evalTerm (TLabelOf t)
  = TLabelOf (evalTerm t)
evalTerm (TApp (TLam x t) tx)
  = subst (x,tx) t
evalTerm (TApp t tx)
  = TApp (evalTerm t) tx
evalTerm v
  = v

eval (Pg lc (TUnlabel (TLabeled l t)))
  = Pg (l ⊔ lc) (TReturn t)

eval (Pg lc (TToLabeled (TLabel l) t))
  | Pg lc' (TLIO t') ← eval* (Pg lc t)
  , lc  ⊑ l, lc' ⊑ l
  = Pg lc (TReturn (TLabeled l t'))
  | otherwise
  = Pg lc TException

eval (Pg lc t)
 = Pg lc (evalTerm t)

eval PgHole
 = PgHole

eval* :: Label l ⇒ Program l → Program l
eval* PgHole
  = PgHole
eval* (Pg lc (TLIO t))
  = Pg lc (TLIO t)
eval* p
  = eval* (eval p)

subst :: Eq l ⇒ (Int, Term l)
      → Term l → Term l
subst = ...
```

Fig. 6. Operational semantics of λ_{LIO}.

In our mechanization we extended pure terms with lattice label operations, branches, lists, and inductive fixpoints; we omit them here for space reasons.

Monadic terms. Monadic terms are evaluated under a state that captures the current label. Bind (`TBind`) and return (`TReturn`) are the standard monadic operations, that respectively propagate and return the current state. The current label is accessed with the `TGetLabel` term and the monadic term `TLIO` wraps monadic values, i.e., computations that cannot be further evaluated. The term `TTLabel lt t` labels the term `t` with the label term `lt` and dually the term `TUnlabel t` unlabels the labeled term `t`. An exception (`TException`) is thrown if a policy is violated. Finally, the term `TToLabeled tl t` locally raises the current label to `tl` to evaluate the monadic term `t`, dropping it again when the computation completes.

3.2.2 Semantics. Figure 6 summarizes the operational semantics of λ_{LIO} as three main functions, (i) `eval` evaluates monadic terms taking into account the current label of the program, (ii) `evalTerm` evaluates pure terms, and (iii) `eval*` is the transitive closure of `eval`.

Program evaluation. The bind of two terms `t1` and `t2` fully evaluates `t1` into a monadic value, using evaluation's transitive closure `eval*`. The result is passed to `t2`. The returned program uses the label of the evaluation of `t1`, which is safe since evaluation only increases the current label. In the definition of evaluation, we use Haskell's guard syntax `Pg lc' (TLIO t1') ← eval* (Pg lc t1)` to denote that evaluation of bind only occurs when `eval* (Pg lc t1)` returns a program whose term is a monadic value `TLIO`. Using refinement types, we prove that assuming that programs cannot diverge and are well-typed (i.e., `t1` is a monadic term), `eval* (Pg lc t1)` always returns a program with a monadic value, so evaluation of bind always succeeds. Evaluation of the `TReturn` term simply returns a monadic value and evaluation of `TGetLabel` returns the current label. Evaluation of `TTLabel (TLabel l) t` returns the term `t` labeled with `l`, when the current label can flow to `l`, otherwise it returns an exception. Dually, unlabeling `TLabeled l t` returns the term `t` with the current label joined with `l`. The term `ToLabeled (TLabel l) t` under current label `lc` fully evaluates the term `t` into a monadic value `t'` with returned label `lc'`. If both the current and returned labels can flow into `l`, then evaluation returns the term `t` labeled with the returned label `lc'`, while the current label remains the same. That is, evaluation of `t` can arbitrarily raise the label, since its result is labeled under `l`. Otherwise, an exception is thrown. The rest of the terms are pure, and their evaluation rules are given below. Finally, evaluation of a hole is an identity.

Term evaluation. Evaluation of the term `TLabelOf t` returns the label of `t`, if `t` is a labeled term; otherwise it propagates evaluation until `t` is evaluated to a labeled term. Evaluation of application uses the standard call-by-name semantics. The definition of substitution is standard and omitted. The rest of the pure terms are either values or a variable, whose evaluation is defined to be the identity. We define `eval*` to be the transitive closure of `eval`. That is, `eval*` repeats evaluation until a monadic value is reached.

3.3 Noninterference

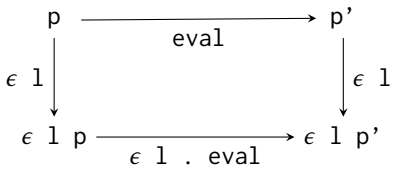
Now we prove noninterference for λ_{LIO}. Noninterference holds when the *low view* of a program is preserved by its evaluation. This low view is characterized by an *erasure* function, which removes program elements whose security label is higher than the adversary's label, replacing them with a "hole." Two versions of the program given possibly different secrets will start with the same low view, and if the program is noninterfering, they will end with the same low view. We prove nointerference of λ_{LIO} by employing a simulation lemma, in the style of Li and Zdancewic [2010]; Russo et al. [2008]; Stefan et al. [2011]. We use refinement types to express this lemma and the property of noninterference, and rely on Liquid Haskell to certify our proof.

3.3.1 Erasure. The functions ϵ and ϵ`Term` erase the sensitive data of programs and terms, *resp.*

```
ϵ     :: Label l ⇒ l → Program l → Program l
ϵTerm :: Label l ⇒ l → Term l    → Term
ϵTerm l (TLabeled l1 t)
  | l1 ⊑ l    = TLabeled l1 (ϵTerm l t)
  | otherwise = TLabeled l1 THole
ϵTerm l (TTLabel (TLabel l1) t)
 | l1 ⊑ l     = TTLabel (TLabel l1) (ϵTerm l t)
 | otherwise  = TTLabel (TLabel l1) THole
...

ϵ l (Pg lc t)
 | lc ⊑ l    = Pg lc (ϵTerm l t)
 | otherwise = PgHole
ϵ _ PgHole   = PgHole
```

The term erasure function ϵ`Term l` replaces terms labeled with a label `l1` with a hole, if `l1` cannot flow into the erasure label `l`. Similarly, term erasure preemptively replaces the term `t` in `TTLabel (TLabel l1) t` with a hole when `l1` cannot flow into the erasure label `l`, since evaluation




```
measure terminates :: Program l → Bool

simulation :: Label l ⇒ l:l
  → p:{Program l | terminates p }
  → { ϵ l (eval (ϵ l p)) = ϵ l (eval p) }
```

Fig. 7. Simulation between `eval` and `ϵ l . eval`.

will lead to a labeled term. For the remaining terms, erasure is a homomorphism. Program erasure with label `l` of a program with current label `lc` erases the term of the program, if `lc` can flow into `l`; otherwise it returns a program hole hiding from the attacker all the program configuration (i.e., both the term and the current label). Erasure of a program hole is an identity.

3.3.2 Simulation. In Figure 7 we state that for every label `l`, `eval` and `ϵ l . eval` form a simulation. That is, evaluation of a program `p` and evaluation of its erased version `ϵ l p` cannot be distinguished after erasure. We prove this property by induction on the input program term.

Termination. Simulation (and later, noninterference) is termination-insensitive: it is defined only for executions that terminate, as indicated by the `terminates` predicate. (λ_{LIO} includes untyped lambda calculus, so λ_{LIO} programs are not strongly normalizing.) This is necessary because, for soundness, Liquid Haskell disallows non-terminating functions, like `eval`, from being lifted into refinement types. To lift `eval` in the logic we constrained it to only be called on terminating programs. To do so, we defined two logical, uninterpreted functions.

```
measure terminates :: Program l → Bool
measure evalSteps  :: Program l → Int
```

We use a refinement-type precondition to prescribe that `eval` is only called on programs `p` that satisfy the `terminates` predicate, and prove termination of `eval` by checking that the steps of evaluation (`evalSteps p`) are decreasing at each recursive call.

```
eval :: Label l ⇒ p:{Program l | terminates p} → Program l / [evalSteps p]
```

While the functions `terminates` and `evalSteps` cannot be defined as Haskell functions, we can instead *axiomatize* properties that are true under the assumption of termination. In particular,

- if a program terminates, so do its subprograms, and
- if a program terminates, its evaluation steps are strictly smaller than those of its subprograms.

To express these properties, we define axioms involving these functions in refinements for each source program construct. For instance, the following assumption (encoded as a Haskell function) handles bind terms:

```
assume evalStepsBindAxiom :: lc:l → db:DB l → t1:Term l
 → t2:{Term l | terminates (Pg lc db (TBind t1 t2)) } →
{  (evalSteps (Pg lc db t1) < evalSteps (Pg lc db (TBind t1 t2)))
  && (0 <= evalSteps (Pg lc db t1))
  && (terminates (Pg lc db t1))} }
evalStepsBindAxiom _ _ _ _ = ()
```

Here, `evalStepsBindAxiom` encodes that if the program `Pg lc db (TBind t1 t2)` terminates, then so does `Pg lc db t1` with fewer evaluation steps. This assumption is required to prove simulation in the inductive case of the `TBind`, since we need to

- apply the simulation lemma for the `Pg lc db t1` program, thus we need to know that it terminates; and
- prove that the induction is well founded, which we do by proving that the evaluation step counts of each subprogram are a decreasing natural number.

3.3.3 Noninterference. The noninterference theorem states that if two terminating λ_{LIO} programs `p1` and `p2` are equal after erasure with label `l`, then their evaluation is also equal after erasure with label `l`. As with simulation, noninterference is termination insensitive—potentially diverging programs could violate noninterference.

We express the noninterference theorem as a refinement type.

```
nonInterference :: Label l ⇒ l:l
 → p1:{Program l  | terminates p1 } → p2:{Program l  | terminates p2 }
 → { ϵ l p1 == ϵ l p2 } → { ϵ l (eval p1) == ϵ l (eval p2) }
```

The proof proceeds by simple rewriting using the simulation property at each input program and the low equivalence precondition.

```
nonInterference l p1 p2 lowEquivalent
  =   ϵ l (eval p1)       ? simulation l p1
  ==. ϵ l (eval (ϵ l p1)) ? lowEquivalent
  ==. ϵ l (eval (ϵ l p2)) ? simulation l p2
  ==. ϵ l (eval p2)
  *** QED
```

The body of `nonInterference` starts from the left hand side of the equality and, using equational reasoning and invocation of the `lowEquivalent` and the `simulation` theorem on the input programs `p1` and `p2`, reaches the right hand side of the equality. As explained in 3.1 the proof combinator `x ? p` returns its first argument and extends the SMT environment with the knowledge of the theorem `p`. The proof combinator `x ==. y = y` equates its two arguments and returns the second argument to continue the equational steps. Finally, `x *** QED = ()` casts its first argument into `unit`, so that the equational proof returns a `unit` type.

4 LABEL-BASED SECURITY FOR DATABASE OPERATIONS

In this section we extend λ_{LIO} with support for databases with label-based policies. We call the extended calculus λ_{LWeb}. In § 4.1, we define a database that stores rows with three values: a key, a first field with a static label, and a second field whose label is a function of the first field. This simplification of the full generality of LWeb's implementation (which permits any field to be a label) captures the key idea that fields can serve as labels for other fields in the same row, and fields that act as labels must be labeled as well. In § 4.2 we define operations to insert, select, delete, and update the database. For each of these operations, in § 4.3 we define a monadic term that respects the database policies. Finally in § 4.4 we define erasure of the database and prove noninterference.

4.1 Database Definition

Figure 8 contains Haskell definitions used to express the semantics of database operations in λ_{LWeb}. Rather than having concrete syntax (e.g., as in Figure 3) for database definitions, in our formalization we assume that databases are defined directly in the semantic model.

A database `DB l` maps names (`Name`) to tables (`Table l`). A table consists of a policy (`TPolicy l`) and a list of rows (`[Row l]`). Each row contains three terms: the key and two values. We limit values that can be stored in the database to basic terms such as unit, integers, label values, etc. This restriction is expressed by predicate `isDBValue`. Labeled terms are not permitted—labels of stored

```
type DB l     = [(Name, Table l)]
type Name     = String
data Table l  = Table {tpolicy :: TPolicy l, tRows :: [Row l]}
data Row l    = Row {rKey :: Term l, rVal1 :: DBTerm l, rVal2 :: DBTerm l }
type DBTerm l = {t:Term l | isDBValue t }
data TPolicy l  = TPolicy { tpTableLabel  :: l , tpFresh :: Int
                          , tpLabelField1 :: {l1:l | l1 ⊑ tpTableLabel }
                          , tpLabelField2 :: Term l → l }
```

Fig. 8. Definition of λ_{LWeb} database

data are specified using the table policy. In § 4.4 we define erasure of the database to replace values with holes, thus `isDBValue` should be true for holes too, but is false for any other term.

```
isDBValue :: Term l → Bool       isDBValue TUnit      = True
isDBValue THole    = True        isDBValue (TLabel _) = True
isDBValue (TInt _) = True        isDBValue _          = False
```

We define the refinement type alias `DBTerm` to be terms refined to satisfy the `isDBValue` predicate and define rows to contain values of type `DBTerm`.

Table policy. The table policy `TPolicy l` defines the security policy for a table. The field `tpTableLabel` is the label required to access the length of the table. The field `tpLabelField1` is the label required to access the first value stored in each row of the table. This label is the same for each row and it is refined to flow into the `tpTableLabel`. The field `tpLabelField2` defines the label of the second value stored in a row as a function of the first. Finally, the field `tpFresh` is used to provide a unique term key for each row. The term key is an integer term that is increased at each row insertion.

Helper functions. For each field of `TPolicy`, we define a function that given a table accesses its respective policy field.

```
labelT   t = tpTableLabel  (tpolicy t)    labelF2  t v = tpLabelField2 (tpolicy t) v
labelF1  t = tpLabelField1 (tpolicy t)    freshKey t   = tpFresh  (tpolicy t)
```

We use the indexing function `db!!n` to lookup the table named `n` in the database.

```
(!!) :: DB l → Name → Maybe (Table l)
```

4.2 Querying the Database

Predicates. We use predicates to query database rows. In the `LWeb` implementation, predicates are written in a domain-specific query language, called `lsql`, which the `LWeb` compiler can analyze. Rather than formalizing that query language in λ_{LWeb}, we model predicates abstractly using the following datatype:

```
data Pred = Pred { pVal :: Bool , pArity :: { i:Int | 0 <= i <= 2 } }
```

Here, `pVal` represents the outcome of evaluating the predicate on an arbitrary row, and `pArity` represents which of the row's fields were examined during evaluation. That is, a `pArity` value of `0`, `1`, or `2`, denotes whether the predicate depends on (i.e., computes over) none, the first, or both fields of a row, respectively.

Then, we define a logical *uninterpreted* function `evalPredicate` that evaluates the predicate for some argument of type `a`:

```
measure evalPredicate :: Pred → a → Bool
```

```
data Program l =
 Pg { pLabel :: l, pDB :: DB l, pTerm :: Term l } | PgHole { pDB :: DB l }

data Term l = ...
 | TInsert Name (Term l) (Term l) | TSelect Name Pred
 | TDelete Name Pred              | TUpdate Name Pred (Term l)
```

Fig. 9. Extension of programs and terms with a database.

We define a Haskell (executable) function `evalPredicate` and use an axiom to connect it with the synonymous logical uninterpreted function [Vazou et al. 2018]:

```
assume evalPredicate :: p:Pred → x:a → {v:Bool | v == evalPredicate p x }
evalPredicate p x = pVal p
```

This way, even though the Haskell function `evalPredicate p x` returns a constant boolean ignoring its argument `x`, the Liquid Haskell model assumes that it behaves as an uninterpreted function that does depend on the `x` argument (with dependencies assumed by the `pArity` definition).

Primitive queries. It is straightforward to define primitive operators that manipulate the database but do not perform IFC checks. We define operators to insert, delete, select, and update databases.

```
(+=) :: db:DB l → n:Name → r:Row l → DB l                                  -- insert
(?=) :: db:DB l → n:Name → p:Pred  → Term l                                -- select
(-=) :: db:DB l → n:Name → p:Pred  → DB l                                  -- delete
(:=) :: db:DB l → n:Name → p:Pred  → v1:DBTerm l → v2:DBTerm l → DB l      -- update
```

Insert: `db += n r` inserts the row `r` in the `n` table in the database and increases `n`'s unique field.
Select: `db ?= n p` selects all the rows of the `n` table that satisfy the predicate `p` as a list of labeled terms.
Delete: `db -= n p` deletes all the rows of the `n` table that satisfy the predicate `p`.
Update: `db := n p v1 v2` updates each row with key `k` of the `n` table that satisfies the predicate `p` with `Row k v1 v2`.

Next we extend the monadic programs of § 3 with database operations to define monadic query operators that enforce the table and field policies.

4.3 Monadic Database Queries

4.3.1 Syntax. Figure 9 defines λ_{LWeb}'s syntax as an extension of λ_{LIO}. Programs are extended to carry the state of the database. Erasure of a program at an observation level `l` leads to a `PgHole` that now carries a database erased at level `l`. Erasure is defined in § 4.4; here we note that preserving the database at program erasure is required since even though the result of the program is erased, its effects on the database persist. For instance, when evaluating `TBind t1 t2` the effects of `t1` on the database affect computing `t2`.

Terms are extended with monadic database queries. `TInsert n (TLabeled l1 v1) (TLabeled l2 v2)` inserts into the table `n` database values `v1` and `v2` labeled with `l1` and `l2`, respectively. `TSelect n p` selects the rows of the table `n` that satisfy the predicate `p`. `TDelete n p` deletes the rows of the table `n` that satisfy the predicate `p`. Finally, `TUpdate n p (TLabeled l1 v1) (TLabeled l2 v2)` updates the fields for each row of table `n` that satisfies the predicate `p` to be `v1` and `v2`, where the database values `v1` and `v2` are labeled with `l1` and `l2`, respectively.

4.3.2 Semantics. Figure 10 defines the operational semantics for the monadic database queries in λ_{LWeb}. Before we explain the evaluation rules, note that both insert and update attempt to insert a labeled value `TLabeled li vi` in the database, thus `vi` should be a value, and unlabeled, i.e., satisfy the `isDBValue` predicate.[3] In the LWeb implementation we use Haskell's type system to enforce this requirement. In λ_{LWeb}, we capture this property in a predicate ς that constrains labeled values in insert and update to be database values:

```
ς :: Program l → Bool
ς (Pg _ _ t) = ςTerm t

ςTerm :: Term l → Bool
ςTerm (TInsert _ (TLabeled _ v1) (TLabeled _ v1))   = isDBValue v1 && isDBValue v2
ςTerm (TUpdate _ _ (TLabeled _ v1) (TLabeled _ v1)) = isDBValue v1 && isDBValue v2
...
```

We specify that `eval` is only called on well-structured programs, i.e., those that satisfy ς. For terms other than insert and update, well-structuredness is homomorphically defined. Restricting well-structuredness to permit only database values, as opposed to terms that eventually evaluate to database values, was done to reduce the number of cases for the proof, but does not remove any conceptual realism.

Insert. Insert attempts to insert a row with values `v1` and `v2`, labeled with `l1` and `l2` respectively, in the table `n`. To perform the insertion we check that

(1) the table named `n` exists in the database, as table `t`.
(2) `l1` can flow into the label of the first field of `t`, since the value `v1` labeled with `l1` will write to the first field of the table.
(3) `l2` can flow into the label of the second field of `t`, as potentially determined by the first field `v1` (i.e., per `labelF2 t v1`).
(4) the current label `l` can flow to the label of the table, since insert changes the length of the table.

If all these checks succeed, we compute a fresh key `k = freshKey t`, insert the row `Row k v1 v2` into the table `n`, and return the key. If any of the checks fail we return an exception and leave the database unchanged.

Either way, we raise the current label `l` by joining it with `l1`. This is because checking `l2` $\sqsubseteq$ `labelF2 t v1` requires examining `v1`, which has label `l1`. That this check succeeds can be discerned by whether the key is returned; if the check fails an exception is thrown, potentially leaking information about `v1`. This subtle point was revealed by the formalization: Our original implementation failed to raise the current label properly.

Select. Select only checks that the table `n` exists in the database, returning an exception if it does not. If the table `n` is found as the table `t`, then we return the term `db ?= n p` that contains a list of all rows of `t` that satisfy the predicate `p`, leaving the database unchanged. The current label is raised to include the label of the table `labelT t` since on a trivially true predicate, all the table is returned, thus the size of the table can leak. We raise the current label with the label of the predicate `p` on the table `t` that intuitively permits reading all the values of `t` that the predicate `p` depends on. We define the function `labelPred p t` that computes the label of the predicate `p` on the table `t`.

```
labelPred  :: (Label l) ⇒ Pred → Table l → l
labelPred p (Table tp rs)
```

[3]We could allow inserting unlabeled terms, the label for which is just the current label. Explicit labeling is strictly more general.

```
eval :: Label l ⇒ i:{ Program l | ς i && terminates i} → {o:Program l | ς o }
```

```
eval (Pg l db (TInsert n t1 t2)
 | TLabeled l1 v1 ← t1
 , TLabeled l2 v2 ← t2
 , Just t ← db!!n, l1 ⊑ labelF1 t
 , l2 ⊑ labelF2 t v1, l ⊑ labelT t
 = let k = freshKey t
       r = Row k v1 v2 in
   Pg (l ⊔ l1) (db += n r) (TReturn k)

eval (Pg l db (TInsert n t1 t2))
 | TLabeled l1 v1 ← t1
 , TLabeled l2 v2 ← t2
 = Pg (l ⊔ l1) db (TReturn TException)

eval (Pg l db (TDelete n p))
 | Just t ← db!!n
 , l ⊔ labelPred p t ⊑ labelT t
 = let l' = l ⊔ labelRead p t in
   Pg l' (db -= n p) (TReturn TUnit)

eval (Pg l db (TDelete n p))
 | Just t ← db!!n
 = let l' = l ⊔ labelRead p t in
   Pg l' db (TReturn TException)
 | otherwise
 = Pg l  db (TReturn TException)
```

```
eval (Pg l db (TSelect n p))
 | Just t ← db!!n
 = let l' = l ⊔ labelT t
              ⊔ labelPred p t in
   Pg l' db (TReturn (db ?= n p))

eval (Pg l db (TSelect n p))
 = Pg l  db (TReturn TException)

eval (Pg l db (TUpdate n p t1 t2)
 | TLabeled l1 v1 ← t1
 , TLabeled l2 v2 ← t2
 , Just t ← db!!n
 , l ⊔ l1 ⊔ labelPred p t ⊑ labelF1 t
 , l ⊔ l2 ⊔ labelPred p t ⊑ labelF2 t v1
 = let l' = l ⊔ l1 ⊔ labelRead p t ⊔ labelT t
   in Pg l' (db := n p v1 v2) (TReturn TUnit)

eval (Pg l db (TUpdate n p t1 t2)
 | TLabeled l1 v1 ← t1
 , TLabeled l2 v2 ← t2
 , Just t ← db!!n
 = let l' = l ⊔ l1 ⊔ labelRead p t ⊔ labelT t
   in Pg l' db (TReturn TException)
 | otherwise
 = Pg l  db (TReturn TException)
```

Fig. 10. Evaluation of monadic database terms.

```
  | pArity p == 2 = foldl (⊔) (labelF1 tp) [labelF2 tp v1 | Row _ v1 _ ← rs]
  | pArity p == 1 = labelF1 tp
  | otherwise     = ⊥
```

If the predicate `p` depends on both fields, then its predicate is the join of the label of the first field and all the labels of the second fields. If `p` only depends on the first field, then the label of the predicate `p` is the label of the first field. Otherwise, `p` depends on no fields and its predicate is ⊥.

Note that the primitive selection operator `db ?= n p` returns labeled terms protected by the labels returned by the `labelF1` and `labelF2` functions. Since terms are labeled, select does not need to raise the current label to protect values that the predicate `p` does not read.

Delete. Deletion checks that the table named `n` exists in the database as `t` and that the current label joined with the label of the predicate `p` on the table `t` can flow into the label of the table `t`, since delete changes the size of the table. If both checks succeed, then database rows are properly deleted. The current label is raised with the "read label" of the predicate `p` on the table `t` that intuitively gives permission to read the label of the predicate `p` on the same table. The function `labelRead p t` computes the read label of the predicate `p` on the table `t` to be the label required

to read `labelPredRow p t`, i.e., equal to the label of the first field, if the predicate depends on the second field and bottom otherwise.

```
labelRead :: (Label l) ⇒ Pred → Table l → l
labelRead p t = if pArity p == 2 then labelF1 t else ⊥
```

Note that `labelRead p t` always flows into `labelPred p t`, thus the current label is implicitly raised to this read label. When the runtime checks of `delete` fail we return an exception and the database is not changed. If the table `n` was found in the database, the current label is raised, even in the case of failure, since the label of the predicate was read.

Update. Updating a table `n` with values `v1` and `v2` on a predicate `p` can be seen as a select-delete-insert operation. But, since the length of the table is not changing, the check that the current label can flow to the label of the table is omitted. Concretely, update checks that

(1) the table named `n` exists in the database, as table `t`,
(2) `l ⊔ l1 ⊔ labelPred p t` can flow into the label of the first field of `t`, since the value `v1` labeled with `l1` will write on the first field of the table and whether this write is done or not depends on the label of the predicate `p` as a hole,
(3) `l ⊔ l2 ⊔ labelPred p t` can flow into the label of the second field of `t` when the first field is `v1`.

If these checks succeed, then unit is returned, the database it updated, and the current label is raised to all the labels of values read during the check, i.e., `l1 ⊔ labelF1 t`. If the checks fail then we return an exception and the database is not updated.

In both cases, the current label is raised by joining with the table label, i.e., `l' = ... ⊔ labelT t`. This is because the last check depends on whether the table is empty or not, and its success can be discerned: if it succeeds, then unit is returned. Interestingly, our original implementation failed to update the current label in this manner. Doing so seemed intuitively unnecessary because an update does not change the table length.

4.4 Noninterference

As in § 3 to prove noninterference we prove the simulation between `eval` and ϵ `l . eval` for λ_{LWeb} programs. Figure 11 extends erasure to programs and databases. Erasure of programs is similar to § 3 but now we also erase the database. Erasure of a database recursively erases all tables. Erasure of a table removes all of its rows if the label of the table cannot flow into the erasing label, thus hiding the size of the table. Otherwise, it recursively erases each row. Erasure of a row respects the dynamic labels stored in the containing table's policy. Erasure of a row replaces *both* fields with holes if the label of the first field cannot flow into the erasing label, since the label of the second field is not visible. If the label of the second field cannot flow into the erasing label, it replaces only the second field with a hole. Otherwise, it erases both fields.

With this definition of erasure, we prove the simulation between `eval` and ϵ `l . eval`, and with this, noninterference. The refinement properties in the database definition of fig. 8 are critical in the proof, as explained below.

Well-structured programs. The simulation proof assumes that the input program is well-structured, i.e., satisfies the predicate ς as defined in § 4.3.2, or equivalently evaluation only inserts values that satisfy the `isDBValue` property. To relax this assumption, an alternative approach could be to check this property at runtime, just before insertion of the values. But, this would break simulation: `TInsert n (TLabeled l1 v1) t` will fail if `v1` is not a database value, but its erased version can succeed if `v1` is erased to a hole (when `l1` cannot flow into the erase label). Thus, the

```
ϵ      :: (Label l) ⇒ l → Program l → Program l
ϵDB    :: (Label l) ⇒ l → DB l → DB l
ϵTable :: (Label l) ⇒ l → Table l → Table l
ϵRow   :: (Label l) ⇒ l → TPolicy l → Row l → Row l

ϵ l (PgHole db)
  = PgHole (ϵDB l db)
ϵ l (Pg lc db t)
 | ¬ (lc ⊑ l)
 = PgHole (ϵDB l db)
 | otherwise
 = Pg lc (ϵDB l db) (ϵTerm l t)

ϵDB l []
 = []
ϵDB l ((n,t):db)
 = (n,ϵTable l):ϵDB l db

ϵTable l (Table tp rs) | ¬ (tpTableLabel tp ⊑ l)
 = Table tp []
ϵTable l (Table tp rs)
 = Table tp (map (ϵRow l tp) rs)

ϵRow l tp (Row k v1 v2)
 | ¬ (tpLabelField1 tp ⊑ l)
 = Row k THole THole
 | ¬ (tpLabelField2 tp v1 ⊑ l)
 = Row k (ϵTerm l v1) THole
 | otherwise
 = Row k (ϵTerm l v1) (ϵTerm l v2)
```

Fig. 11. Erasure of programs and databases.

`isDBValue` property cannot be checked before insertion and should be assumed by evaluation. In the implementation this safety check is enforced by Haskell's type system.

Database values. Simulation of the delete operation requires that values stored in the database must have identity erasure, e.g., cannot be labeled terms. Thus, we prove that all terms that satisfy `isDBValue` also have erasure identity. We do this by stating the property as a refinement on term erasure itself.

```
ϵTerm :: Label l ⇒ l → i:Term l → {o:Term l | isDBValue i ⇒ isDBValue o }
```

In the delete proof, each time a database term is erased, the proof identity `ϵTerm l v == v` is immediately available.

Note on refinements. The type `DBTerm l` is a type alias for `Term l` with the attached refinement that the term is a database value. A `DBTerm l` *does not carry* an actual proof that it is a database value. Instead, the refinement type that the term satisfies the `isDBValue` property is statically verified during type checking. As a consequence, comparison of two `DBTerms` does not require proof comparison. At the same time, verification can use the `isDBValue` property. For instance, when opening a row `Row k v1 v2`, we know that `isDBValue v1` and by the type of term erasure, we know that for each label `l`, `ϵTerm l v1 == v1`.

5 LIQUID HASKELL FOR METATHEORY

Liquid Haskell was originally developed to support lightweight program verification (e.g., out-of-bounds indexing). The formalization of `LWeb` in Liquid Haskell, presented in § 3 and § 4, was made possible by recent extensions to support general theorem proving [Vazou et al. 2018]. Our proof of noninterference was a challenging test of this new support, and constitutes the first advanced metatheoretical result mechanized in Liquid Haskell.[4]

The trusted computing base (TCB) of any Liquid Haskell proof relies on the correct implementation of several parts. In particular, we trust that

[4] https://github.com/plum-umd/lmonad-meta

(1) the GHC compiler correctly desugars the Haskell code to the core language of Liquid Haskell,
(2) Liquid Haskell correctly generates the verification conditions for the core language, and
(3) the SMT solver correctly discharges the verification conditions.

We worked on the noninterference proof, on and off, for 10 months. The proof consists of 5,447 lines of code and requires about 5 hours to be checked. For this proof in particular, we (naturally) trust all of our semantic definitions, and also two explicit assumptions, notably the axiomatization of termination and modeling of predicates. These were discussed respectively in § 3.3.2 and § 4.2.

Carrying out the proof had a clear benefit: As mentioned in § 4.3, we uncovered two bugs in our implementation. In both cases, `LWeb` was examining sensitive data when carrying out a security check, but failed to raise the current label with the label of that data. Failure of the mechanized proof to go through exposed these bugs.

The rest of this section summarizes what we view as the (current) advantages and disadvantages of using Liquid Haskell as a theorem prover compared to other alternatives (e.g., Coq and F-star [Swamy et al. 2016]), expanding on a prior assessment [Vazou et al. 2017].

5.1 Advantages

As a theorem proving environment, Liquid Haskell offers several advantages.

General purpose programming language. The Liquid Haskell-based formal development is, in essence, a Haskell program. All formal definitions (presented in § 3 and § 4) and proof terms (e.g., illustrated in § 3.3) are Haskell code. Refinement types define lemmas and theorems, referring to these definitions. In fact, some formal definitions (e.g., the `Label` class definition) were taken directly from the implementation. The first author of the paper and main developer of the proof is a Haskell programmer, thus he did not need to learn a new programming language (e.g., Coq) to develop the formal proof. During development we used Haskell's existing development tools, including the build system, test frameworks, and deployment support (e.g., Travis integration).

SMT automation. Liquid Haskell, like Dafny [Leino 2010] and F-star [Swamy et al. 2016], uses an SMT solver to automate parts of the proof, especially the ones that make use of boolean reasoning, reducing the need for manual case splitting. For example, proving simulation for row updates normally proceeds by case splitting on the relative can-flow-to relation between four labels. The SMT automates the case splitting.

Semantic termination checking. To prove termination of a recursive function in Liquid Haskell it suffices to declare a non negative integer value that is decreasing at each recursive call. The `LWeb` proof was greatly simplified by the semantic termination checker. In a previous Coq LIO proof [Stefan et al. 2017], the evaluation relation apparently requires an explicit *fuel* argument to count the number of evaluation steps, since the evaluation function (the equivalent to that in fig. 6) does not necessarily terminate. In our proof, termination of evaluation was axiomatized (per § 3.3.2), which in practice meant that the evaluation steps were counted only in the logic and not in the definition of the evaluation function.

Intrinsic and *extrinsic verification.* The Liquid Haskell proving style allows us to conveniently switch between (manual) extrinsic and (SMT automated) intrinsic verification. Most of the `LWeb` proof is extrinsic, i.e., functions are defined to state and prove theorems about the model. In few cases, intrinsic specifications are used to ease the proof. For instance, the refinement type specification of ϵ`Term`, as described in 4.4, intrinsically specifies that erasure of `isDBValue` terms returns terms that also satisfy the `isDBValue` predicate. This property is automatically proven by the SMT without cluttering the executable portion of the definition with proof terms.

5.2 Disadvantages

On the other hand, Liquid Haskell has room to improve as a theorem proving environment, especially compared to advanced theorem provers like Coq.

Unpredictable verification time. The first and main disadvantage is the unpredictability of verification times, which owe to the invocation of an SMT solver. One issue we ran across during the development of our proof is that internal transformations performed by `ghc` can cause massive blowups. This is because Liquid Haskell analyzes Haskell's intermediate code (`CoreSyn`), not the original source. As an example of the problem, using `|x,y` instead of the logical `| x && y` in function guards leads to much slower verification times. While the two alternatives have exactly the same semantics, the first case leads to exponential expansion of the intermediate code.

Lack of tactics. Liquid Haskell currently provides no tactic support, which could simplify proof scripts. For example, we often had to systematically invoke label laws (fig. 4) in our proofs, whereas a proof tactic to do so automatically could greatly simplify these cases.

General purpose programming language. Liquid Haskell, developed for light-weight verification of Haskell programs, lacks various features in verification-specific systems, such as Coq. For example, Liquid Haskell provides only experimental support for curried, higher-order functions, which means that one has to inline higher order functions, like `map`, `fold`, and `lookup`. There is also no interactive proof environment or (substantial) proof libraries.

In sum, our `LWeb` proof shows that Liquid Haskell can be used for sophisticated theorem proving. We are optimistic that current disadvantages can be addressed in future work.

6 IMPLEMENTATION

`LWeb` has been available online since 2016 and consists of 2,664 lines of Haskell code.[5] It depends on our base `LMonad` package that implements the `LMonadT` monad transformer and consists of 345 lines of code.[6] `LWeb` also imports `Yesod`, a well established, external Haskell library for type-safe, web applications. This section explains how the implementation extends the formalization, and then discusses the trusted computing base.

6.1 Extensions

The `LWeb` implementation generalizes the formalization of Sections 3 and 4 in several ways.

Clearance label. The implementation supports a *clearance* label, described in § 2.1. Intuitively, the clearance label limits how high the current label can be raised. If the current label ever exceeds the clearance label, an exception is thrown. This label is not needed to enforce noninterference, but serves as an optimization, cutting off transactions whose current label rises to the point that they are doomed to fail. Adding checks to handle the clearance was straightforward.

Full tables and expressive queries. As first illustrated in § 2.3, tables may have more than two columns, and a column's label can be determined by other various fields in the same row. The labels of such *dependency fields* must be constant, i.e., not determined by another field, and flow into the table label (which also must be constant). A consequence of this rule is that a field's label cannot depend on itself. Finally, values stored in tables instantiate `Yesod`'s `PersistField` type class. The implementation uses only the predefined instances including `Text`, `Bool`, `Int` but critically, does not define a `PersistField` for labeled values. `LWeb` enforces these invariants at compile time

[5] https://github.com/jprider63/lmonad-yesod
[6] https://github.com/jprider63/lmonad

via Haskell type checking and when preprocessing table definitions. LWeb rewrites queries to add labels to queried results.

We have implemented database operations beyond those given in § 4, to be more in line with typical database support. Some of these operations are simple variations of the ones presented. For example, LWeb allows for variations of `update` that only update specific fields (not whole rows). LWeb implements these basic queries by wrapping Persistent [Snoyman 2018], Yesod's database library, with the derived IFC checks. To support more advanced queries, LWeb defines an SQL-like domain-specific language called `lsql`. `lsql` allows users to write expressive SQL queries that include inner joins, outer joins, **`where`** clauses, orderings, limits, and offsets. Haskell expressions can be included in queries using anti-quotation. At compile-time, LWeb parses `lsql` queries using quasi-quotation and Template Haskell [Sheard and Jones 2002]. It rewrites the queries to be run using Esqueleto [Esqueleto 2018], a Haskell library that supports advanced database queries. As part of this rewriting, LWeb inserts IFC checks for queries based on the user-defined database policies. We show several examples of `lsql` queries in § 7.

Optimizations. Sometimes a label against which to perform a check is derived from data stored in every row. Retrieving every row is especially costly when the query itself would retrieve only a fraction of them. Therefore, when possible we compute an upper bound for such a label. In particular, if a field is fully constrained by a query's predicate, we use the field's constrained value to compute any dependent labels. When a field is not fully constrained, we conservatively set dependent labels to $\top$. Suppose we wish to query the `Friends` table from fig. 3, retrieving all rows such that `user1 == 'Alice'` and `date < '2000-01-01'`. The confidentiality portion of `user1`'s label is $\bot$, but that portion of `date`'s is computed from `user1` $\sqcap$ `user2`. Since `user1` is always `'Alice'` we know the computed label is $\bigsqcup_l$ `Alice` $\sqcap\, l$ for all values `user2` $= l$ in the database. In this case, we can bound l as $\top$, and thus use label `Alice`, since it is equivalent to `Alice` $\sqcap\, \top$. While this bound is technically conservative, in practice we find it makes policy sense. In this example, if the `user2` field can truly vary arbitrarily then $\bigsqcup_l\ l$ will approach $\top$.

Declassification. LWeb supports forms of *declassification* [Sabelfeld and Sands 2009] for cases when the IFC lattice ordering needs to be selectively relaxed. These should be used sparingly (and are, in our BIBIFI case study), as they form part of the trusted computing base, discussed below.

Row ordering. As a final point, we note that our formalization models a database as a list of rows; insertion (via +=) simply appends to the list, regardless of the contents of a row. As such, *row ordering* does not depend on the database's contents and thus reveals nothing about them (it is governed only by the table label). In the implementation, advanced operations may specify an ordering. LWeb prevents leaks in this situation by raising the current label with the label of fields used for sorting. If a query does not specify an ordering, LWeb takes no specific steps. However, ordering on rows is undefined in SQL, so a backend database could choose to order them by their contents, and thus potentially leak information in a query's results. In our experience with PostgreSQL, default row ordering depends on when values are written and is independent of the data in the table.

6.2 Trusted Computing Base

A key advantage of LWeb is that by largely shifting security checks from the application into the LWeb IFC framework, we can shrink an application's trusted computing base (TCB). In particular, for an application that uses LWeb, the locus of trust is on LWeb itself, which is made (more) trustworthy by our mechanized noninterference proof. A few parts of the application must be trusted, nevertheless.

First, all of the policy specifications are trusted. The policy includes the labels on the various tables and the labels on data read/written from I/O channels. Specifying the latter requires writing

```
UserInfo
  user       UserId
  school     Text  <Const Admin ⊓ Field user, Field user>
  age        Int   <Const Admin ⊓ Field user, Field user>
  experience Int   <Const Admin ⊓ Field user, Field user>
```

Fig. 14. Table `UserInfo` contains additional BIBIFI user information.

some trusted code to interpret data going in or out. For example, in a multi-user application like BIBIFI, code performing authentication on a particular channel must be trusted (§ 7.2).

Second, any uses of declassification are trusted, as they constitute local modifications to policy. One kind of declassification can occur selectively on in-application data [Sabelfeld and Myers 2003]. We give an example in § 7.4. Another kind of declassification is to relax some security checks during database updates. The update query imposes strong runtime checks, e.g., that the label of the predicate should flow into the updated fields as formalized in § 4. LWeb provides an unsound update alternative (called `updateDeclassifyTCB`) that ignores this specific check.

7 THE BIBIFI CASE STUDY

As a real world web application of `LWeb`, we present *Build it, Break it, Fix it* (BIBIFI), a security-oriented programming contest [Ruef et al. 2016] hosted at https://builditbreakit.org. The contest consists of three rounds. At the outset, the organizers publish the specification for some software that has particular security goals. During the first round, teams implement software to this specification, aiming for it to be both fast and secure. In the second round, teams find as many breaks as possible in the implementations submitted by other teams. During the final round, teams attempt to fix the identified problems in their submissions.

7.1 BIBIFI Labels

BIBIFI labels include all entities that operate in the system. The `Principal` data type, defined in fig. 12, encodes all such entities, including the system itself, the administrator, users, teams, and judges. Each of these entities is treated as a security level. For instance a policy can encode that data written by a user with id `5`, can get protected at the security level of this specific user, so that only he or she can read this data. A more flexible policy encodes that the system administrator can read data written by each user. To encode such policies, we use disjunction category labels (`DCLabel`) [Stefan et al. 2012] to create a security lattice out of our `Principal`s. In fig. 12 we define `BBFLabel` as the `DCLabel Principal` data type that tracks the security level of values as they flow throughout the web application and database.

```
data Principal
 = PSys | PAdmin | PUser UserId
 | PTeam  TeamId | PJudge JudgeId

type BBFLabel = DCLabel Principal
```

Fig. 12. BIBIFI labels.

7.2 Users and Authentication

Users' personal information is stored in the BIBIFI database. Figure 14 shows the `User` table with the basics: a user's account id, email address, and whether they have administrator privileges. The label for the `email` field refers to `Id` in its label: This is a shorthand for the key of the present table. The label says

```
User
  account Text
  email Text   <Const Admin ⊓ Id, Id>
  admin Bool   <⊥, Const Admin>
```

Fig. 13. Basic BIBIFI `User` table.

```
Announcement <⊥, Const Admin>          Team                        TeamMember
  title   Text <⊥ Const Admin>           name    Text                team TeamId
  content Text <⊥, Const Admin>          contest ContestId           user UserId
```

Fig. 15. Definition of Announcement, Team, and TeamMember tables and their policies.

that a user can read and write their emails, while the administrator can read every user's email. The label for the admin field declares that it may be written by the administrator and read by anyone.

Additional private information is stored in the UserInfo table, shown in Figure 14, including a user's school, age, and professional experience. The user field of this table is a foreign key to the User table, as indicated by its type UserId (see § 2.2). Each of the remaining fields is protected by this field, in part: users can read and write their own information while administrators can read any users' information.

The current label is set by the code trusted to perform authentication. If a user is not logged in, the current label is set to <⊥, ⊤>: the confidentiality label is the upper bound on data read so far (i.e., none, so ⊥), and the integrity label is the level of least trust (i.e., ⊤) for writing data. After authenticating, most users will have the label <⊥, PUser userId>, thus lowering the integrity part (thus increasing the level of trust) to the user itself. Users who are also administrators will have current label lowered further to <⊥, PUser userId ⊓ PAdmin>. This is shown in the following code snippet. It determines the logged in user via requireAuth, and then adds administrator privileges if the user has them (per userAdmin).

```
(Entity userId user) ← requireAuth
let userLabel = dcIntegritySingleton (PrincipalUser userId)
lowerLabelTCB $ if userAdmin user
                  then userLabel ⊓ dcIntegritySingleton PrincipalAdmin
                  else userLabel
```

The clearance is also set using trusted functions during authentication. For example, for an administrator it would be <PUser userId ⊔ PAdmin, ⊤>.

7.3 Opening the Contest

To start a contest, administrators write announcements that include information like instructions and problem specifications. It is important that only administrators can post these announcements. Announcements are stored in the database, and their (simplified) table definition is shown in fig. 15. The Announcement table has two Text fields corresponding to an announcement's title and content. Only administrators can author announcements.

An earlier version BIBIFI relied on manual access control checks rather than monadic LMonad enforcement of security. The old version had a security bug: it failed to check that the current user was an administrator when posting a new announcement. Here is a snippet of the old code.

```
postAddAnnouncementR :: Handler Html
postAddAnnouncementR = do
    ((res, widget), enctype) ← runFormPost postForm
    case res of ...
        FormSuccess (FormData title markdown) → do
            runDB (insert (Announcement title markdown))
            redirect AdminAnnouncementsR
```

This function parses POST data and inserts a new announcement. The user is never authenticated, so anyone can post new announcements and potentially deface the website. In the IFC version of the website, the database insertion fails for unauthorized or unauthenticated users as the integrity part of the current label is not sufficiently trusted (the label does not flow into `PAdmin`).

7.4 Teams and Declassification

To participate in a contest, a user must join a team. The teams and their members are stored in the eponymous tables of fig. 15. Teams serve as another principal in the BIBIFI system and BIBIFI defines a TCB function that appropriately authenticates team members similarly to users (§ 7.2), authorizing a team member to read and write data labeled with their team.

BIBIFI uses declassification (as discussed in 6.2) to allow team members to send email messages to their team. The policy on the email field of the `User` table states that only the user or an administrator can read the email address, so BIBIFI cannot give a user's email address to a teammate. Instead, the function `sendEmailToTeam` below sends the email on the teammate's behalf using declassification.

```
sendEmailToTeam :: TeamId → Email → LHandler ()
sendEmailToTeam tId email = do
  protectedEmails ← runDB [lsql| pselect User.email from User inner join
      TeamMember on TeamMember.user == User.id where TeamMember.team == #{tId}
      |]
  mapM_ (\protectedEmail → do
      address ← declassifyTCB protectedEmail
      sendEmail address email
    ) protectedEmails
```

The function `sendEmailToTeam`'s parameters are the team identifier and an email return address. It queries the database for the (labeled) email addresses of the team's members, using `lsql` (see § 2.3 and § 6.1). The `sendEmailToTeam` function maps over each address, declassifying it via `declassifyTCB`, so that the message can be sent to the address. The `declassifyTCB` function takes a labeled value and extracts its raw value, *ignoring label restrictions*. This is an unsafe operation that breaks noninterference, so the programmer must be careful with its use. Here for example, the function is careful not to reveal the email address to the sender but only use it to send the email.

7.5 Breaks and Advanced Queries

During the second round of the BIBIFI context, teams submit breaks, i.e., test cases that attack another team's submission. After a break is pushed to a registered git repository, BIBIFI's backend infrastructure uploads it to a virtual machine and tests whether the attack succeeds. Results are stored in the `BreakSubmission` table of fig. 16, which has fields for the attacking team, the target team, and the (boolean) result of the attack. The integrity label for the result field is `PSys` since only the backend system can grade an attack. The confidentiality label is `PAdmin ⊓ PTeam attackerId ⊓ PTeam targetId` since administrators, the attacker team, and the target team can see the result of an attack.

BIBIFI has an administration page that lists all break submissions next to which team was attacked. This page's contents are retrieved via the following inner join.

```
runDB $ [lsql| select BreakSubmission.★, Team.name from BreakSubmission inner join
  Team on BreakSubmission.target == Team.id where Team.contest == #{contestId} order
  by BreakSubmission.id desc |]
```

```
BreakSubmission
  attacker TeamId <⊥, Const Sys>
  target   TeamId <⊥, Const Sys>
  result   Bool   <Const Admin ⊓ Field attacker ⊓ Field target, Const Sys>
```

Fig. 16. Definition of `BreakSubmission` table and its policy.

Table 1. Latency comparison between the **Vanilla** and **LWeb** implementations of the BIBIFI application. The mean, standard deviation, and tail latency in milliseconds over 1,000 trials are presented. In addition, the response size in kilobytes and the overhead of LWeb are shown.

Handler	Verb	Vanilla Latency			LWeb Latency			Size (kB)	Overhead
		Mean (ms)	SD (ms)	Tail (ms)	Mean (ms)	SD (ms)	Tail (ms)		
/announcements	GET	4.646	1.215	16	5.529	1.367	20	18.639	19.01%
/announcement/update	POST	9.810	2.600	54	11.395	3.054	52	0.706	16.16%
/profile	GET	2.116	0.512	6	2.167	0.550	6	7.595	2.41%
/buildsubmissions	GET	6.364	1.251	17	7.441	1.706	22	14.434	16.92%
/buildsubmission	GET	28.633	2.772	52	30.570	3.477	75	9.231	6.76%
/breaksubmissions	GET	41.758	7.826	81	49.218	11.679	90	60.044	17.86%
/breaksubmission	GET	4.070	0.538	9	4.923	0.509	9	6.116	20.96%

This query performs a join over the `BreakSubmission` and `Team` tables, aligning rows where the target team equals the team's identifier. In addition, it filters rows to the specified contest identifier and orders results by the break submission identifiers.

8 EXPERIMENTAL EVALUATION

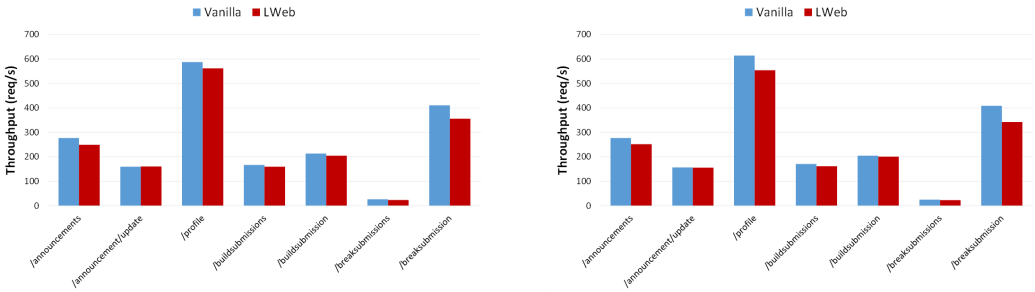
To evaluate `LWeb` we compare the BIBIFI implementation that uses `LMonad` with our initial BIBIFI implementation that manually checked security policies via access control. We call this initial version the *vanilla implementation*. Transitioning from the vanilla to the `LWeb` implementation reduced the trusted computing base (TCB) but imposed a modest runtime overhead.

8.1 Trusted Computing Base of BIBIFI

The implementation of the BIBIFI application is currently 11,529 lines of Haskell code. 80 of these lines invoke trusted functions (for authentication or declassification, see § 6.2). `LWeb`'s library is 3,009 lines of trusted code. The vanilla implementation is several years old, with 7,367 LOC; there is no IFC mechanism so the whole codebase is trusted. Switching from the vanilla to the `LWeb` implementation only added 151 LOC. The size of the TCB is now 21% of the codebase; considering only the code of the BIBIFI web application (and not `LWeb` too), 1% of the code is trusted.

8.2 Running Time Overhead

We measured the query latency, i.e., the response time (in milliseconds) of HTTP requests, for both the `LWeb` and the vanilla implementation. Measurements were performed over `localhost` and we ran 100 requests to warm up. We present the mean, standard deviation, and tail latency over 1,000 trials, as well as the response size (in kilobytes) and the overhead of `LWeb` over the vanilla implementation. Table 1 summarizes this comparison. The server used for benchmarking runs Ubuntu 16.04 with two Intel(R) Xeon(R) E5-2630 2.60GHz CPUs and 64GB of RAM. PostgreSQL 9.5.13 is run locally as the database backend. We used ApacheBench to perform the measurements with a concurrency level of one. Here is a sample invocation of `ab`:



(a) Concurrency level of 16.

(b) Concurrency level of 32.

Fig. 17. Throughput (req/s) of the **Vanilla** and **LWeb** versions of the BIBIFI application.

```
ab -g profile_lweb.gp -n 1000 -T "application/x-www-form-urlencoded; charset=UTF-8" -c 1
    -C _SESSION=... http://127.0.0.1:4000/profile
```

Most of the requests are GET requests that display contest announcements, retrieve a user's profile with personal information, get the list of a team's submissions, and view the results of a specific submission. One POST request is measured that updates the contents of an announcement. Cookies and CSRF tokens were explicitly defined so that a user was logged into the site, and the user had sufficient permissions for all of the pages.

To evaluate LWeb's impact on the throughput of web applications, we conduct similar measurements except we rerun ab with concurrency levels of 16 and 32. The rest of the experimental setup matches that of the latency benchmark, including number of requests, hardware, and handlers. Figure 17 shows the number of requests per second for each version of the BIBIFI web application across the various handlers.

Most of the handlers show modest overhead between the vanilla and LWeb versions of the website. We measure LWeb's overhead to range from 2% to 21%, which comes from the IFC checks that LWeb makes for every database query and the state monad transformer that tracks the current label and clearance label. In practice, this overhead results in a few milliseconds of delay in response times. In most situations, this is a reasonable price to pay in order the reduce the size of the TCB and increase confidence that the web application properly enforces the user defined security policies.

9 RELATED WORK

LWeb provides end-to-end information flow control (IFC) security for webapps. Its design aims to provide highly expressive policies and queries in a way that does not compromise security, and adds little overhead to transaction processing, in both space and time. This section compares LWeb to prior work, arguing that it occupies a unique, and favorable, spot in the design space.

Information flow control. LWeb is part of a long line of work on using lattice-ordered, label-based IFC to enforce security policies in software [Bell and LaPadula 1973; Denning 1976; Sabelfeld and Myers 2003]. Enforcement can occur either *statically* at compile-time, e.g., as part of type checking [Lourenço and Caires 2014, 2015; Myers 1999; Pottier and Simonet 2003] or a static analysis [Arzt et al. 2014; Hammer and Snelting 2009; Johnson et al. 2015], or *dynamically* at run-time, e.g., via source-to-source rewriting [Chudnov and Naumann 2015; Hedin et al. 2016] or library/run-time support [Roy et al. 2009; Stefan et al. 2011; Tromer and Schuster 2016]. Dynamic approaches often work by rewriting a program to insert the needed checks and/or by relying on support from the hardware, operating system, or run-time. Closely related to IFC, *taint tracking*

controls *data flows* through the program, rather than overall influence (which includes effects on *control flow*, i.e., *implicit* flows). Taint tracking can avoid the false positives of IFC, which often overapproximates control channels, but will also miss security violations [King et al. 2008].

LWeb builds on the LIO framework [Stefan et al. 2011], which is a dynamic approach to enforcing IFC that takes advantage of Haskell's static types to help localize checks to I/O boundaries. LIO's *current label* and *clearance label* draw inspiration from work on Mandatory Access Control (MAC) operating systems [Bell and LaPadula 1973], including Asbestos [Efstathopoulos et al. 2005], HiStar [Zeldovich et al. 2006], and Flume [Krohn et al. 2007]. The baseline LIO approach has been extended in several interesting ways [Buiras et al. 2017, 2015; Russo 2015; Waye et al. 2017], including to other languages [Heule et al. 2015].

The proof of security in the original LIO (without use of a database) has been partially mechanized in Coq [Stefan et al. 2017], while the derivative MAC library [Russo 2015] has been mechanized in Agda [Vassena and Russo 2016]. The MAC mechanization considers concurrency, which ours does not. Ours is the first mechanization to use an SMT-based verifier (Liquid Haskell).

IFC for database-using web applications. Several prior works apply IFC to web applications. FlowWatcher [Muthukumaran et al. 2015] enforces information flow policies within a web proxy, which provides the benefit that applications need not be retrofitted, but limits the granularity of policies it can enforce.

SeLINQ [Schoepe et al. 2014] is a static IFC system for F# programs that access a database via language-integrated queries (LINQ). SIF [Chong et al. 2007b] uses Jif [Myers 1999] to enforce static IFC-based protection for web servlets, while Swift [Chong et al. 2007a] also allows client-side (Javascript) code. Unlike LWeb, these systems permit only statically determined database policies, not ones with dynamic labels (e.g., stored in the database). The latter two lack language support for database manipulation, though a back-end database can be made accessible by wrapping it with a Jif signature (which we imagine would require an SeLINQ-style static policy).

UrFlow [Chlipala 2010] performs static analysis to prove that information flow policies are properly enforced. These policies are expressed as SQL queries over protected data and known information. Static analysis-based proofs about queries and flows impose no run-time overhead. But static analysis can be overapproximate, rejecting correct programs. Dynamic enforcement schemes do not have this issue, and LWeb's LIO-based approach imposes little run-time overhead.

SELinks [Corcoran et al. 2009] enforces security policies for web applications, including ones resembling the field-dependent policies we have in LWeb. To improve performance, security policy checks were offloaded to the database as stored procedures; LWeb could benefit from a similar optimization. SELinks was originally based on a formalism called Fable [Swamy et al. 2008] in which one could encode IFC policies, but this encoding was too onerous for practical use, and not present in SELinks, which was limited to access control policies. Qapla [Mehta et al. 2017] also supports rich policies, but like SELinks these focus on access control, and so may fail to plug leaks of protected data via other server state.

Jacqueline [Yang et al. 2016] uses faceted information flow control [Austin and Flanagan 2012] to implement policy-agnostic security [Austin et al. 2013; Yang et al. 2012] in web applications. Like LWeb, they have formalized and proved a noninterference property (but not mechanized it). Unlike LWeb that enforces IFC using the underlying LIO monad, Jacqueline at runtime explicitly keeps track of the secret and public views of sensitive values. While expressive, this approach can be expensive in both space and time: results of computations on sensitive values have up to 1.75× slower running times, and require more memory. Latencies for Django and Jacqueline are around 160ms for typical requests to their benchmark application.

The system most closely related to LWeb is Hails [Giffin et al. 2017, 2012], which aims to enforce information flow-oriented policies in web applications. Hails is also based on LIO, and is particularly interested in confining third-party extensions (written in Safe Haskell [Terei et al. 2012]). In Hails, individual record fields can have policies determined by other data in the database, as determined by a general Haskell function provided by the programmer. Thus, Hails policies can encode LWeb policies, and more; e.g., data in one table can be used to determine labels for data in another table. Evaluating the policy function during query processing is potentially expensive. That said, according to their benchmarks, the throughput of database writes of Hails is 2× faster than Ruby Sinatra, comparable to Apache PHP, and 6× slow than Java Jetty. They did not measure Hails' overhead, e.g., by measuring the performance difference with and without policy checks.

There are several important differences between LWeb and Hails. First, LWeb builds on top of a mature, popular web framework (Yesod). Extracting LIO into LMonad makes it easy for LWeb to evolve as Yesod evolves. As such, LWeb can benefit from Yesod's optimized code, bugfixes, etc. Second, LWeb's `lsql` query language is highly expressive, whereas (as far as we can tell) Hails uses a simpler query language targeting MongoDB where predicates can only depend on the document key. Third, there is no formal argument (and little informal argument) that Hails' policy checks ensure a high-level security property. The ability to run arbitrary code to determine policies seems potentially risky (e.g., if there are mutually interacting policy functions), and there seems to be nothing like our database invariants that are needed for noninterference. Our mechanized formalization proved important: value-oriented policies (where one field's label depends on another field) were tricky to get right (per § 5).

Finally, IFDB [Schultz and Liskov 2013] defines an approach to integrating information flow tracking in an application and a database. Like Hails and LWeb, the application tracks a current "contamination level," like LIO's current label, that reflects data it has read. In IFDB, one can specify per-row policies using secrecy and integrity labels, but not policies per field. Labels are stored as separate, per-row metadata, implemented by changing the back-end DBMS. Declassification is permitted within trusted code blocks. Performance overhead for HTTP request latencies was similar to LWeb, at about 24%. Compared to IFDB, LWeb does not require any PSQL/database modifications; can support per-field, updatable labels; and can treat existing fields as labels, rather than requiring the establishment of a separate (often redundant) field just for the label. IFDB also lacks a clear argument for security, and has no formalization. Once again, we found such a formalization particularly useful for revealing bugs.

10 CONCLUSION

We presented LWeb, a information-flow security enforcement mechanism for Haskell web applications. LWeb combines Yesod with LMonad, a generalization of the LIO library. LWeb performs label-based policy checks and protects database values with dynamic labels, which can depend on the values stored in the database. We formalized LWeb (as λ_{LWeb}) and used Liquid Haskell to prove termination-insensitive noninterference. Our proof uncovered two noninterference violations in the implementation. We used LWeb to build the web site of the *Build it, Break it, Fix it* security-oriented programming contest, and found it could support rich policies and queries. Compared to manually checking security policies, LWeb impose a modest runtime overhead between 2% to 21% but reduces the trusted code base to 1% of the application code, and 21% overall (when counting LWeb too).

ACKNOWLEDGMENTS

We would like to thank Alejandro Russo and the anonymous reviewers for helpful comments on a draft of this paper. This work was supported in part by the National Science Foundation under grant CNS-1801545 and by DARPA under contract FA8750-16-C-0022.

REFERENCES

Steven Arzt, Siegfried Rasthofer, Christian Fritz, Eric Bodden, Alexandre Bartel, Jacques Klein, Yves Le Traon, Damien Octeau, and Patrick McDaniel. 2014. FlowDroid: Precise Context, Flow, Field, Object-sensitive and Lifecycle-aware Taint Analysis for Android Apps. In *PLDI*.

Thomas H. Austin and Cormac Flanagan. 2012. Multiple Facets for Dynamic Information Flow. In *POPL*.

Thomas H. Austin, Jean Yang, Cormac Flanagan, and Armando Solar-Lezama. 2013. Faceted execution of policy-agnostic programs. In *PLAS*.

D. Elliott Bell and Leonard J. LaPadula. 1973. Secure Computer Systems: Mathematical Foundations. *MITRE Technical Report 2547* 1 (1973).

Pablo Buiras, Joachim Breitner, and Alejandro Russo. 2017. Securing concurrent lazy programs against information leakage. In *CSF*.

Pablo Buiras, Dimitrios Vytiniotis, and Alejandro Russo. 2015. HLIO: Mixing Static and Dynamic Typing for Information-flow Control in Haskell. In *ICFP*.

Adam Chlipala. 2010. Static Checking of Dynamically-varying Security Policies in Database-backed Applications. In *OSDI*.

Stephen Chong, Jed Liu, Andrew C. Myers, Xin Qi, K. Vikram, Lantian Zheng, and Xin Zheng. 2007a. Secure Web Applications via Automatic Partitioning. In *SOSP*.

Stephen Chong, K. Vikram, and Andrew C. Myers. 2007b. SIF: Enforcing Confidentiality and Integrity in Web Applications. In *USENIX Security Symposium*.

Andrey Chudnov and David A. Naumann. 2015. Inlined Information Flow Monitoring for JavaScript. In *CCS*.

Brian J. Corcoran, Nikhil Swamy, and Michael Hicks. 2009. Cross-tier, Label-based Security Enforcement for Web Applications. In *SIGMOD*.

Dorothy E. Denning. 1976. A Lattice Model of Secure Information Flow. *Commun. ACM* 19, 5 (1976).

Petros Efstathopoulos, Maxwell Krohn, Steve VanDeBogart, Cliff Frey, David Ziegler, Eddie Kohler, David Mazières, Frans Kaashoek, and Robert Morris. 2005. Labels and Event Processes in the Asbestos Operating System. In *SOSP*.

Esqueleto 2018. Esqueleto: Type-safe EDSL for SQL queries on persistent backends. https://github.com/bitemyapp/esqueleto/blob/master/docs/blog_post_2012_08_06.

Daniel Giffin, Amit Levy, Deian Stefan, David Terei, David Mazieres, John Mitchell, and Alejandro Russo. 2017. Hails: Protecting data privacy in untrusted web applications. *Journal of Computer Security* 25, 4 (2017).

Daniel B. Giffin, Amit Levy, Deian Stefan, David Terei, David Mazières, John C. Mitchell, and Alejandro Russo. 2012. Hails: Protecting Data Privacy in Untrusted Web Applications. In *OSDI*.

Joseph A. Goguen and José Meseguer. 1982. Security Policies and Security Models. In *S&P*.

Christian Hammer and Gregor Snelting. 2009. Flow-sensitive, Context-sensitive, and Object-sensitive Information Flow Control Based on Program Dependence Graphs. *International Journal of Information Security* 8, 6 (2009).

Daniel Hedin, Luciano Bello, and Andrei Sabelfeld. 2016. Information-flow security for JavaScript and its APIs. *Journal of Computer Security* 24, 2 (2016).

Stefan Heule, Deian Stefan, Edward Z. Yang, John C. Mitchell, and Alejandro Russo. 2015. IFC Inside: Retrofitting Languages with Dynamic Information Flow Control. In *POST*.

Andrew Johnson, Lucas Waye, Scott Moore, and Stephen Chong. 2015. Exploring and Enforcing Security Guarantees via Program Dependence Graphs. In *PLDI*.

Dave King, Boniface Hicks, Michael Hicks, and Trent Jaeger. 2008. Implicit Flows: Can't live with 'em, can't live without 'em. In *International Conference on Information Systems Security (ICISS) (Lecture Notes in Computer Science)*, R. Sekar and Arun K. Pujari (Eds.), Vol. 5352.

Maxwell Krohn, Alexander Yip, Micah Brodsky, Natan Cliffer, M. Frans Kaashoek, Eddie Kohler, and Robert Morris. 2007. Information Flow Control for Standard OS Abstractions. In *SOSP*.

K. Rustan M. Leino. 2010. Dafny: An Automatic Program Verifier for Functional Correctness. In *LPAR*.

Peng Li and Steve Zdancewic. 2010. Arrows for Secure Information Flow. *Theoretical Computer Science* 411, 19 (2010).

Luísa Lourenço and Luís Caires. 2014. Information Flow Analysis for Valued-Indexed Data Security Compartments. In *International Symposium on Trustworthy Global Computing - Volume 8358*.

Luísa Lourenço and Luís Caires. 2015. Dependent Information Flow Types. In *POPL*.

Aastha Mehta, Eslam Elnikety, Katura Harvey, Deepak Garg, and Peter Druschel. 2017. Qapla: Policy compliance for database-backed systems. In *USENIX Security Symposium*.

Divya Muthukumaran, Dan O'Keeffe, Christian Priebe, David Eyers, Brian Shand, and Peter Pietzuch. 2015. FlowWatcher: Defending Against Data Disclosure Vulnerabilities in Web Applications. In *CCS*.

Andrew C. Myers. 1999. JFlow: Practical Mostly-static Information Flow Control. In *POPL*.

François Pottier and Vincent Simonet. 2003. Information Flow Inference for ML. *ACM Trans. Program. Lang. Syst.* 25, 1 (2003).

Indrajit Roy, Donald E. Porter, Michael D. Bond, Kathryn S. McKinley, and Emmett Witchel. 2009. Laminar: Practical Fine-grained Decentralized Information Flow Control. In *PLDI*.

Andrew Ruef, Michael Hicks, James Parker, Dave Levin, Michelle L. Mazurek, and Piotr Mardziel. 2016. Build It, Break It, Fix It: Contesting Secure Development. In *CCS*.

Alejandro Russo. 2015. Functional Pearl: Two Can Keep a Secret, if One of Them Uses Haskell. In *ICFP*.

Alejandro Russo, Koen Claessen, and John Hughes. 2008. A library for light-weight information-flow security in haskell. In *Haskell Symposium*.

Andrei Sabelfeld and Andrew C. Myers. 2003. Language-based information-flow security. *Selected Areas in Communications, IEEE Journal on* 21, 1 (2003).

Andrei Sabelfeld and David Sands. 2009. Declassification: Dimensions and Principles. *Journal of Computer Security* 17, 5 (2009).

Daniel Schoepe, Daniel Hedin, and Andrei Sabelfeld. 2014. SeLINQ: Tracking Information Across Application-database Boundaries. In *ICFP*.

David Schultz and Barbara Liskov. 2013. IFDB: Decentralized Information Flow Control for Databases. In *EuroSys*.

Tim Sheard and Simon Peyton Jones. 2002. Template Meta-programming for Haskell. In *Haskell Workshop*.

Michael Snoyman. 2018. Yesod Web Framework for Haskell. http://www.yesodweb.com/.

Deian Stefan, Alejandro Russo, David Mazières, and John C. Mitchell. 2012. Disjunction Category Labels. In *NordSec*.

Deian Stefan, Alejandro Russo, David Mazières, and John C. Mitchell. 2017. Flexible Dynamic Information Flow Control in the Presence of Exceptions. *Journal of Functional Programming* 27, E5 (2017).

Deian Stefan, Alejandro Russo, John C. Mitchell, and David Mazières. 2011. Flexible Dynamic Information Flow Control in Haskell. In *Haskell Symposium*.

Nikhil Swamy, Brian Corcoran, and Michael Hicks. 2008. Fable: A Language for Enforcing User-defined Security Policies. In *S&P*.

Nikhil Swamy, Catalin Hritcu, Chantal Keller, Aseem Rastogi, Antoine Delignat-Lavaud, Simon Forest, Karthikeyan Bhargavan, Cédric Fournet, Pierre-Yves Strub, Markulf Kohlweiss, Jean-Karim Zinzindohoué, and Santiago Zanella-Béguelin. 2016. Dependent Types and Multi-Monadic Effects in F*. In *POPL*.

David Terei, Simon Marlow, Simon Peyton Jones, and David Mazières. 2012. Safe Haskell. In *Haskell Symposium*.

Eran Tromer and Roei Schuster. 2016. DroidDisintegrator: Intra-Application Information Flow Control in Android Apps. In *ASIA CCS*.

Marco Vassena and Alejandro Russo. 2016. On Formalizing Information-Flow Control Libraries. In *PLAS*.

Niki Vazou, Leonidas Lampropoulos, and Jeff Polakow. 2017. A Tale of Two Provers: Verifying Monoidal String Matching in Liquid Haskell and Coq. In *Haskell Symposium*.

Niki Vazou, Eric L. Seidel, Ranjit Jhala, Dimitrios Vytiniotis, and Simon L. Peyton Jones. 2014. Refinement Types for Haskell. In *ICFP*.

Niki Vazou, Anish Tondwalkar, Vikraman Choudhury, Ryan G. Scott, Ryan R. Newton, Philip Wadler, and Ranjit Jhala. 2018. Refinement Reflection: Complete Verification with SMT. *PACMPL* 2, POPL (2018).

Lucas Waye, Pablo Buiras, Owen Arden, Alejandro Russo, and Stephen Chong. 2017. Cryptographically Secure Information Flow Control on Key-Value Stores. In *CCS*.

Jean Yang, Travis Hance, Thomas H. Austin, Armando Solar-Lezama, Cormac Flanagan, and Stephen Chong. 2016. Precise, Dynamic Information Flow for Database-backed Applications. In *PLDI*.

Jean Yang, Kuat Yessenov, and Armando Solar-Lezama. 2012. A language for automatically enforcing privacy policies. In *POPL*.

Nickolai Zeldovich, Silas Boyd-Wickizer, Eddie Kohler, and David Mazières. 2006. Making Information Flow Explicit in HiStar. In *OSDI*.