An Axiomatic Approach to Detect Information Leaks in Concurrent Programs

Sandip Ghosal  
Department of Computer Science & Engineering  
Indian Institute of Technology, Bombay  
Mumbai, India  
sandipsmit@gmail.com

R. K. Shyamasundar  
Department of Computer Science & Engineering  
Indian Institute of Technology, Bombay  
Mumbai, India  
shyamasundar@gmail.com

Abstract—Realizing flow security in a concurrent environment is extremely challenging, primarily due to non-deterministic nature of execution. The difficulty is further exacerbated from a security angle if sequential threads disclose control locations through publicly observable statements like print, sleep, delay, etc. Such observations lead to internal and external timing attacks. Inspired by previous works that use classical Hoare style proof systems for establishing correctness of distributed (real-time) programs, in this paper, we describe a method for finding information leaks in concurrent programs through the introduction of leaky assertions at observable program points. Specifying leaky assertions akin to classic assertions, we demonstrate how information leaks can be detected in a concurrent context. To our knowledge, this is the first such work that enables integration of different notions of non-interference used in functional and security context. While the approach is sound and relatively complete in the classic sense, it enables the use of algorithmic techniques that enable programmers to come up with leaky assertions that enable checking for information leaks in sensitive applications.

I. INTRODUCTION

Information flow control (IFC) in sequential and concurrent programs has been an active area of research for over two decades. While researchers have developed specification approaches [1], [2], [3], [4] and enforcement mechanisms for sequential programs [5], [6], [7], [8], [9], [10], [11], flow security in concurrent programs remains a challenge primarily due to covert channels that could leak secret information. Timing channels in concurrent programs are the most commonly known covert channels.

![Fig. 1: (a) Process $T_1$ (left); (b) Process $T_2$ (right)](image)

For instance, often programs encode secret information in such a way that allow an attacker to learn about confiden-
used in ‘Gries-Owicki’ [22]. By showing the program is non-interference, we infer information leaks through the leaky assertions embedded in the specification.

Main contributions of the paper are:

1. Introduction of leaky assertions when embedded in concurrent program specifications, enables detection of information leaks once the proof of components of the concurrent programs are shown to be correct when considered in isolation and are further shown to be interference free.
2. Demonstration of the method of embedding leaky assertions in standard concurrent program specification [22], can be used to assert information leaks, from which the information leaks are detected.
3. Checking for information leaks due to public observers like IO and time sensitive operations (those that could disclose program locations through output statements or observable timing channels), in non-interfering flow secure concurrent programs [27].

Rest of the paper is organized as follows. Section II provides an abstract syntax of the concurrent language. Section III briefly describes the necessary background on correctness of concurrent programs and non-interference. Section IV illustrates our solution followed by a formal description of our approach and proof of soundness in Section V. Section VI compares our approach with other related work in the literature, followed by conclusions in Section VII.

II. LANGUAGE SYNTAX

Before proceeding with the solution, we shall briefly describe the syntax of the language we use.

Let \( s \) ranging over set \( S \) be the set of commands, \( x \) be the set of variables ranging over \( X \), and \( e \) denote an expression. Language primarily comprises basic actions such as \texttt{skip}, assignment, sequence and control statements such as conditional branch, iteration, \texttt{await} as in [22]. In addition, we consider two statements \texttt{print(e)} and \texttt{delay(e)} that serve as observable points from the perspective of security and hence, these statements are often termed public statements. Statement \texttt{print(e)} writes \([e]\), where \([.]\) denotes the classic evaluation of expressions, into the standard output file that is visible to public. On the other hand, \texttt{delay(e)} postpones the current execution for \([e]\) time units. The abstract syntax of the language is given below:

\[
\begin{align*}
  s &::= \texttt{skip} \mid x = e \mid s_1 ; s_2 \mid \text{if } e \text{ then } s_1 \text{ else } s_2 \mid \text{while } e \text{ do } s \mid \text{print}(e) \mid \text{delay}(e) \mid s_1 \parallel s_2 \mid \text{await } b \text{ then } s
\end{align*}
\]

Operator \( s_1 \parallel s_2 \) denotes a parallel (or concurrent) composition of sequential programs \( s_1 \) and \( s_2 \). The \texttt{await} command is a conditional critical region, that turns a statement \( s \) into an indivisible action, is executed in a state that satisfies condition \( b \). The semantics of sequential part of our language essentially follows the classical while-language. Whereas, the semantics for concurrent composition and \texttt{await} statements are in line with the transition rules given in [28].

III. CORRECTNESS OF CONCURRENT PROGRAMS

The seminal work of Hoare [29] provided a basis for establishing axiomatic correctness of imperative programs. Let \( P \) and \( Q \) be assertions referred to as pre-condition and post-condition for program/statement \( S \). Hoare’s triple specification for partial correctness of program \( S \), is denoted \( \{P\}S\{Q\} \) with the following interpretation: if \texttt{pre-condition} \( P \) is true before execution of \( S \), then \texttt{post-condition} \( Q \) holds if and when the execution of \( S \) terminates.

Owicki and Gries [22] generalized Hoare’s axiomatic approach for shared variable concurrency with the introduction of non-interference (functional) of programs. Let \( S_i \) for \( i = 1, \ldots, n \), be sequential programs. \( S_1 \parallel S_2 \parallel \cdots \parallel S_n \) denotes concurrent composition. The effect of concurrent execution is the same as executing each \( S_i \) as an independent program if the programs are disjoint. If the programs are not disjoint, then, it is necessary that programs be interference free. Non-interference requires that assertions used in the proof \( \{P_i\}S_i\{Q_i\} \) for each process remain validated under concurrent execution of the other processes. Let, \( T \) be a statement in program \( S_i \). Then, concurrent execution of \( \{P_i\}S_1\{Q_1\}, \ldots, \{P_n\}S_n\{Q_n\} \) is said to be 

\textit{non-interfering} if \( T \) does not interfere with \( \{P_j\}S_j\{Q_j\} \) where \( i \neq j \).

Formal definition is given below.

\textbf{Definition 1 (Correctness).} Given a proof \( \{P\}S\{Q\} \) and a statement \( T \) with the pre-condition \( \text{pre}(T) \), the statement \( T \) is said to be non-interfering with \( \{P\}S\{Q\} \) if the following two conditions hold:

1. \( \{Q \land \text{pre}(T)\}T\{Q\} \);
2. \( \{\text{pre}(S') \land \text{pre}(T)\}T\{\text{pre}(S')\} \) where \( S' \) is an \texttt{await} or assignment statement within \( S \) but not within an \texttt{await} block.

IV. CAPTURING INFORMATION LEAKS THROUGH OUTPUT

The crux of our approach lies in setting observers from which one can draw inference about information leaks. Observation points allow us to deduce how long a process could wait at a control point relative to another process at some other control point wherein the two processes share information via shared variables (could be sensitive) through synchronization. This is realized through auxiliary variables that are used in the specification of concurrent programs. Lamport [25] illustrates an effective use of auxiliary variables for establishing correctness of real-time programs.

Timing constraints for the waiting time that impose upper or lower bound for an action to occur can be expressed using auxiliary (time) variables. Semantic characterization of time constraints required for mutual exclusion and analysis of bounded waiting time is similar to that shown in [25].

\textbf{A. Illustrative Example}

Let us illustrate that the labelled program \( T_3 \) shown in Figure 2, where labels denote identifiers for statements as well as locations of program control does indeed leak information via public statements. Labels \( l_0 \) and \( l_8 \) are the entry and exit labels respectively. Let us mark the execution time of the
program at specific program locations using auxiliary variable t. Execution of statements l₃ to l₅ is subject to availability of

\[
\begin{align*}
l₀ & : (t = T) \quad \text{print('c')} \\
l₁ & : \text{if } h \text{ then} \\
l₂ & : \text{await } \text{sem} > 0 \text{ then} \\
l₃ & : \text{sem} = \text{sem} - 1; \\
l₄ & : \text{v} = \text{v} + 2; \\
l₅ & : \text{sem} = \text{sem} + 1 \\
\text{else} & \quad \text{skip;} \\
l₇ & : (t = t') \quad \text{print('d')} \\
l₈ & : \\
\end{align*}
\]

Fig. 2: Process T2 (Cf. Figure 1) with auxiliary variable t sem, and hence, involves a delay, say \(\Delta t\). Now, an observer can easily measure the intermediate time difference between execution of public statements \(l₀\) and \(l₇\) and just a comparison with \(\Delta t\) would enable him to deduce that information in \(h\) more deterministic (informally, cardinality of the set of possible values gets reduced from the original set of possibilities) than he could infer otherwise.

\[
\begin{align*}
[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \ll (t' - T) \\
\text{h = 0} & \\
[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \gg (t' - T) \\
\text{h = 1} & \\
\end{align*}
\]

In short, our approach is based on specifying postulates of leaky assertions at expected control points in the program and establish correctness of the program in a standard manner. Once the proof is established, the leaky assertions locations indicate the points of leakage of information; the actual path can be derived using the semantic rules of the language for the program.

Thus, embedding leaky assertions in the specification of functional correctness (Definition 1) of programs leads to a proof system that enables asserting leaks in the programs. Let \(S\) to be an assignment or a non-nested await statement from the thread \(T₁\). Let \(P\) and \(Q\) be pre- and post-conditions of \(S\) respectively. Then, proof of \(\{P\} S \{Q\}\) asserts that the leaky assertions hold when \(T₁\) and \(T₂\) execute concurrently.

1 \(\{Q\} \land \{(\{[at₀] : t = T \land [at₇] : t = t' \land t' - T \ll \Delta t) \implies (h = 0)\} \lor \{(\{[at₀] : t = T \land [at₇] : t = t' \land t' - T \gg \Delta t) \implies (h = 1)\}\}
\]

2 \(\{P\} \land \{([at₀] : t = T \land [at₇] : t = t' \land t' - T \ll \Delta t) \implies (h = 0)\} \lor \{([at₀] : t = T \land [at₇] : t = t' \land t' - T \gg \Delta t) \implies (h = 1)\}\}

3 \{([at₀] : t = T \land [at₇] : t = t' \land t' - T \ll \Delta t) \implies (h = 0)\} \lor \{([at₀] : t = T \land [at₇] : t = t' \land t' - T \gg \Delta t) \implies (h = 1)\}\}

\(\Delta t \iff (h = 1)\land P\)

Note that the leaky assertions successfully capture the execution time inconsistencies for both the cases where semaphore \(\text{sem}\) is locked or released by the process \(T₁\). Therefore, the assertions remain unaffected by the pre-and post-conditions for any statement \(S\) in \(T₁\). As the assertions always hold good, this invariably proves the program as leaky.

Remarks: It is interesting to understand the impact if the program contains a delay statement; suppose the statement \(\text{skip}\) is replaced by \(\text{delay(50)}\) – a delay of 50ms (milliseconds for instance). Now, consider the following cases with respect to the leaky assertions:

1 \(\Delta t \ll 50ms:\
\]

\[
\begin{align*}
\{[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \ll (t' - T) \geq 50ms \\
\text{h = 0} & \\
\{[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \leq (t' - T) \ll 50ms \\
\text{h = 1} & \\
\end{align*}
\]

2 \(\Delta t \gg 50ms:\
\]

\[
\begin{align*}
\{[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \gg (t' - T) \geq 50ms \\
\text{h = 0} & \\
\{[at₀] & : t = T \quad [at₇] : t = t' \quad \Delta t \leq (t' - T) \gg 50ms \\
\text{h = 1} & \\
\end{align*}
\]

3 \(\Delta t \approx 50ms:\
\]

It would be difficult to postulate leaky assertion as an observer cannot easily determine the execution path for \(h = 0\) or \(h = 1\); hence, the program may be considered secure even though it may not be; note that axiomatic proof is relative to the specifications.

V. Formalization of our Approach

In our two-language approach for establishing leaks, if we can show that assertions at various program points imply information leaks, then we would have established that the program is not-secure; in other words, it is leaky. Thus, if we establish that the program is insecure, we would have established a counter-example; on the other hand, if we establish that the program is secure, it would imply that the program is secure relative to the given specification. Below, we establish the first from which the second follows naturally.

Definition 2 (Leaky Assertion). For a given set of assertions \(\langle a₀, \ldots, aₙ \rangle\) at respective control points \(⟨ℓ₀, \ldots, ℓₙ⟩\) of the given program \(S\), using a nonempty set of sensitive (or high) variables \(h₁, \ldots, hₘ\) of the program, is said to be leaky if \(aₜ\) determines the value of at least one of its’ sensitive variables, say \(h_j\).

As discussed earlier, we say that an execution is leaky if the proof of the program does not invalidate any of the leaky assertions.
Definition 3 (Leaky Execution). Given a proof of \( \{P\}S\{Q\} \) (we ignore global invariant as given in [22]), where \( S \) is an assignment or non-nested \texttt{await} statement, and a program \( T \) with at least one output statement with leaky assertion \( A \) as it’s pre-condition, concurrent execution of \( S \) and \( T \) is said to be leaky if the following two conditions hold:

1. \( \{Q \land A\}T\{Q\} \)
2. \( \{P \land A\}T\{P\} \)

Proposition 1. Let \( \{P\}S\{Q\} \) be given and let \( T \) be a statement with pre-condition \( A \), where \( P, Q \) and \( A \) are security assertions. If \( [S \parallel T] \) is leaky then there exists at least one leaky assertion that determines information in at least one secret variable.

Theorem 1 (Soundness). Given the proof of \( \{P\}S\{Q\} \) that has a leaky assertion \( A_i \) at control point \( i \), within program \( S \), then there exists an execution path that leads to truth of \( A_i \).

Proof. Follows straight from the proof of \( \{P\}S\{Q\} \).

Issues of completeness also follow on the same lines.

Application to Flow Security:
Dynamic labelling (DL) algorithms [30] are in use for compile-time flow security certification. For a program performing IO, however, the DL algorithm could also be effectively used for identifying program points for arriving at leaky assertions at plausible control points. DL algorithm binds subjects and objects of a program with the given static or dynamic labels from a security lattice [31] and evaluates the program as per the flow policy that enforces information flow only in upward direction of the lattice. For standard output, a static label \( \text{low} \) is assigned denoting least confidential policy. The output statements controlled by sensitive variables that are labelled with higher security labels would cause an information flow from higher label to low, violating flow policy; thus, the algorithm easily highlights public statements with possible sensitivities. Now, one could visualize the interactions of the various concurrent components that could possibly interfere in the functional sense. Using such an intuition, one could come up with different leaky assertions that would detect information leaks.

The process of finding information leak in programs that perform IO could be automated through the following steps: (i) first use DL algorithm to identify the public statements that could leak information; (ii) capture time-sensitivity for deriving leaky assertions corresponding to those public statements using various analysis tools or oracles; (iii) arrive at possible interaction of components at these public locations, and (iv) apply model checkers to check for certain properties in components assuming possible assertions in other components. We are working on arriving such algorithmic techniques for detecting information leaks.

VI. RELATED WORK
In this section, we briefly highlight some of the important approaches in the literature to prevent timing leaks in concurrent settings. In particular, we focus on the solution that prevents leaking information through publicly observable control points.

Our idea is primarily motivated by the work shown in [22], [32], [33]. The first one presented flow proof rules that use correctness properties given in [22]. The last two approaches have given a semantic interpretation of non-interference using Hoare-like logic. However, the goal of the above approaches is to establish flow security of programs, whereas we aim to find information leaks due to IO statements and time-sensitive operations using simple inference.

Le Guernic [13] first envisages the potential danger in the use of output statements in concurrent programs, and initial solution that motivated us to provide an alternative approach in the form of security assertions. In this approach, the author develops a run-time monitor that works in tandem with a security automaton to decide whether to allow, deny or modify executions of output statements and synchronization commands, particularly those appear in a branch conditioned on sensitive variables. The automaton forbids executions of output statements in this situation. Further, it acquires locks of synchronization command before evaluating such a conditional. For instance, acquiring semaphore \texttt{sem} in the program shown in Figure 1 would prevent from outputting “acdb”. The approach provides a thoughtful solution; nonetheless, it is too restrictive. E.g., in the absence of an output statement, say \texttt{print(‘c’)}, the program is harmless, but as per the above solution executing the program would still be dangerous. The axiomatic approach proposed in this paper removes such restrictions on the program. Our approach enables us to assert flow insecurity in the program methodically through the “Gries-Owicki” conditions for functional correctness, thus overcomes the cases of false alarms.

Other than the above approach, the majority of the mechanisms found in the literature to prevent timing leaks in concurrent programs are based on either (i) developing a sound typing rules to enforce secure-flow properties in concurrent programming languages [19], [21]: (ii) performing source transformation to balance the time variation or eliminate the timing channel [18], [20], [34]; or (iii) developing a run-time monitor to identify potentially dangerous executions that could leak information [14], [35], [36].

VII. CONCLUSIONS AND FUTURE WORK
In this paper, we have proposed an axiomatic proof system using leaky assertions placed at output statement locations. To the best of our knowledge, we are the first to propose an axiomatic approach using approaches for establishing correctness of real-time distributed programs to analyze information leaks in concurrent programs that have observable public locations, leading to the following advantages: (i) it enables to formally prove the presence of information leaks; (ii) aids in integrating the notions of non-interference used in functional and security contexts with respect to complete information flow models like that of Dorothy Denning [31], for shared variable programs.
We are working towards automating the process of detecting information leaks in concurrent programs that exhibit execution progress through public statements. Automation has potential use in flow security certification of concurrent programs to identify information leaks via sensitive public statements. Initial work shows that DL algorithm can be integrated with model checkers for validating/speculating leaky assertions.

REFERENCES

[1] S. Zdancewic and A. C. Myers, “Observational determinism for concurrent program security,” in 16th IEEE Computer Security Foundations Workshop, 2003. Proceedings. IEEE, 2003, pp. 29–43.
[2] G. Boudol and I. Castellani, “Noninterference for concurrent programs and thread systems,” Theoretical Computer Science, vol. 281, no. 1-2, pp. 109–130, 2002.
[3] D. M. Volpino, C. E. Irvine, and G. Smith, “A sound type system for secure flow analysis,” Journal of Computer Security, vol. 4, no. 2/3, pp. 167–188, 1996.
[4] J. A. Goguen and J. Meseguer, “Security policies and security models,” in IEEE Symposium on SP, 1982, pp. 11–11.
[5] P. Buiras, D. Vytiniotis, and A. Russo, “HLIO: Mixing static and dynamic typing for information-flow control in haskell,” in ACM SIGPLAN Notices, vol. 50, no. 9. ACM, 2015, pp. 289–301.
[6] D. Stefan, A. Russo, J. C. Mitchell, and D. Mazières, “Flexible dynamic information flow control in haskell,” in ACM SigPlan Notices, vol. 46, no. 12, 2011, pp. 95–106.
[7] L. Zheng and A. C. Myers, “Dynamic security labels and static information flow control,” International Journal of Information Security, vol. 6, no. 2-3, pp. 67–84, 2007.
[8] C. Hammer and G. Snelting, “Flow-sensitive, context-sensitive, and object-sensitive information flow control based on program dependence graphs,” Int. Journal of Information Security, vol. 8, no. 6, pp. 399–422, 2009.
[9] V. Simonet and I. Rocquencourt, “Flow caml in a nutshell,” in Proc. of 1st APPEMiS-II workshop, 2003, pp. 152–165.
[10] A. C. Myers, L. Zheng, S. Zdancewic, S. Chong, and N. Nystrom, “Tif: Java information flow,” http://www.cs.cornell.edu/jif/, 2001.
[11] D. E. Denning and P. J. Denning, “Certification of programs for secure information flow,” Communications of the ACM, vol. 20, no. 7, pp. 504–513, 1977.
[12] P. Buiras, D. Stefan, and A. Russo, “On dynamic flow-sensitive floating-label systems,” in Proc. of IEEE 27th CSF Symposium, 2014, pp. 65–79.
[13] G. La Guernec, “Automaton-based confidentiality monitoring of concurrent programs,” in 20th IEEE Computer Security Foundations Symposium (CSF’07). IEEE, 2007, pp. 218–232.
[14] M. Vassena, G. Soeller, P. Amidon, M. Chan, J. Renner, and D. Stefan, “Foundations for parallel information flow control runtime systems,” in International Conference on Principles of Security and Trust. Springer, 2019, pp. 1–28.
[15] M. Vassena, J. Breitner, and A. Russo, “Securing concurrent lazy programs against information leakage,” in 2017 IEEE 30th Computer Security Foundations Symposium (CSF). IEEE, 2017, pp. 37–52.
[16] A. Popescu, J. Holzli, and T. Nipkow, “Formal verification of language-based concurrent noninterference,” Journal of Formalized Reasoning, vol. 6, no. 1, pp. 1–30, 2013.
[17] G. Barthe, T. Rezk, A. Russo, and A. Sabelfeld, “Security of multi-threaded programs by compilation,” ACM Transactions on Information and System Security (TISSEC), vol. 13, no. 3, pp. 1–32, 2010.
[18] A. Russo, J. Hughes, D. Naumann, and A. Sabelfeld, “Closing internal timing channels by transformation,” in Annual Asian Computing Science Conference. Springer, 2006, pp. 120–135.
[19] G. Smith et al., “A new type system for secure information flow,” in CSFW, vol. 1, 2001, p. 4.
[20] J. Agat, “Transforming out timing leaks,” in Proceedings of the 27th ACM SIGPLAN-SIGACT symposium on Principles of programming languages, 2000, pp. 40–53.
[21] G. Smith and D. Volpino, “Secure information flow in a multi-threaded imperative language,” in Proceedings of the 25th ACM SIGPLAN-SIGACT symposium on Principles of programming languages. ACM, 1998, pp. 355–364.
[22] S. Owicki and D. Gries, “An axiomatic proof technique for parallel programs i,” Acta informatica, vol. 6, no. 4, pp. 319–340, 1976.
[23] F. B. Schneider, B. Bloom, and K. Marzullo, “Putting time into proof outlines,” in Workshop/School/Symposium of the REX Project (Research and Education in Concurrent Systems). Springer, 1991, pp. 618–639.
[24] L. Lamport, “An assertional correctness proof of a distributed algorithm,” Science of Computer Programming, vol. 2, no. 3, pp. 175–206, 1982.
[25] ———, “Real time is really simple,” Microsoft Research, pp. 2005–30, 2005.
[26] M. Abadi and L. Lamport, “An old-fashioned recipe for real time,” ACM Transactions on Programming Languages and Systems (TOPLAS), vol. 16, no. 5, pp. 1543–1571, 1994.
[27] G. R. Andrews and R. P. Reitman, “An axiomatic approach to information flow in programs,” ACM Transactions on Programming Languages and Systems (TOPLAS), vol. 2, no. 1, pp. 56–76, 1980.
[28] S. Brookes, “Full abstraction for a shared-variable parallel language,” Information and Computation, vol. 127, no. 2, pp. 145–163, 1996.
[29] C. A. R. Hoare, “An axiomatic basis for computer programming,” Communications of the ACM, vol. 12, no. 10, pp. 576–580, 1969.
[30] S. Ghosal, R. K. Shyamansundar, and N. V. N. Kumar, “Static security certification of programs via dynamic labelling,” in Proceedings of the 15th International Joint Conference on e-Business and Telecommunications, ICETE 2018 - Volume 2: SECRIPT, Porto, Portugal, July 26-28, 2018., pp. 400–411.
[31] D. E. Denning, “A lattice model of secure information flow,” CACM, vol. 19, no. 5, pp. 236–243, 1976.
[32] T. Amtsof and A. Banerjee, “Information flow analysis in logical form,” in International Static Analysis Symposium. Springer, 2004, pp. 100–115.
[33] L. Beringer and M. Hofmann, “Secure information flow and program logics,” in 20th IEEE Computer Security Foundations Symposium (CSF’07). IEEE, 2007, pp. 233–248.
[34] A. Russo and A. Sabelfeld, “Security for multithreaded programs under cooperative scheduling,” in International Andrei Ershov Memorial Conference on Perspectives of System Informatics. Springer, 2006, pp. 474–480.
[35] D. Stefan, A. Russo, P. Buiras, A. Levy, J. C. Mitchell, and D. Mazières, “Addressing covert termination and timing channels in concurrent information flow systems,” in ACM SIGPLAN Notices, vol. 47, no. 9, 2012, pp. 201–214.
[36] A. Askarov, S. Chong, and H. Mantel, “Hybrid monitors for concurrent noninterference,” in 2015 IEEE 28th Computer Security Foundations Symposium. IEEE, 2015, pp. 137–151.