Increasing Network Resiliency by Optimally Assigning Diverse Variants to Routing Nodes

Andrew Newell[1], Thomas Tantillo[2], Daniel Obenshain[2], Cristina Nita-Rotaru[1], and Yair Amir[2]
[1]Department of Computer Science, Purdue University
{newella,crisn}@cs.purdue.edu
[2]Department of Computer Science, Johns Hopkins University
{tantillo,dano,yairamir}@cs.jhu.edu

Technical Report CNDS-2013-1 - October 2013
http://www.dsn.jhu.edu
http://projects.cerias.purdue.edu/ds2/intr-tol-clouds.html

***Abstract*—Networks with homogeneous routing nodes are constantly at risk as any vulnerability found against a node could be used to compromise all nodes. Introducing diversity among nodes can be used to address this problem. With few variants, the choice of assignment of variants to nodes is critical to the overall network resiliency.**

We present the Diversity Assignment Problem (DAP), the assignment of variants to nodes in a network, and we show how to compute the optimal solution in medium-size networks. We also present a greedy approximation to DAP that scales well to large networks. Our solution shows that a high level of overall network resiliency can be obtained even from variants that are weak on their own.

We provide two variations of our problem to meet real-world system needs. First, for networks with knowledge of higher-level protocols we offer a technique to create assignments that maximize the needs of a specific application (e.g., Paxos and BFT). Second, for networks with knowledge of the value of traffic between each communicating pair of nodes, we offer a weighted version that can increase resiliency between important communicating pairs while sacrificing resiliency for the less important pairs.

Our assignments are based on assumed compromise probabilities and independence of compromises between different diverse variants. We provide analysis when these assumed probabilities or independence are inaccurate.

I. Introduction

Networks with homogeneous routing nodes are constantly at risk as any vulnerability found against a single routing node could be used to compromise all nodes. Diversity can be employed at various levels on the routing nodes to address this problem by improving resiliency against different classes of attacks. In this work, we base resiliency on the number of surviving client-to-client connections offered by the network when under attack. Diversifying the operating system provides protection against common types of attacks that target operating system vulnerabilities [1]; utilizing multi-variant programming protects against programming vulnerabilities or logical programming errors [2], [3]; using different administrative personnel mitigates social engineering or insider attacks [4]. However, there are only a limited number of operating systems, software versions, and personnel to utilize as diverse variants. So then, how does one assign these limited number of diverse variants to the routing nodes in the network to achieve optimal resiliency?

Initially, we assumed that a random assignment of a few diverse variants would perform well. However, we were surprised to find that a random assignment performs rather poorly, in many cases providing less resiliency than using the best single variant at all routing nodes, and occasionally even less resiliency than using the worst single variant at all routing nodes. Clearly, a better approach is necessary to realize the benefits of diversity.

Our interest in this question arose from constructing a cloud service over a global network of data centers [5]. We needed to have an intrusion-tolerant infrastructure in order to monitor and control the cloud even in the case of sophisticated attacks. While designing intrusion-tolerant protocols for messaging and maintaining consistent state, we realized that without diversity all the nodes could be compromised by a single vulnerability. Inspired by [1], we were especially interested in diversifying the operating system (e.g., Linux, MacOS, and FreeBSD). The additional overhead of managing multiple operating systems within the cloud infrastructure led us to consider only a small number of variants to create diversity.

In this paper, we demonstrate that the way diverse variants are assigned across the network (i.e., which variant is assigned to which routing node) is of utmost importance to the overall network resiliency when the number of variants is smaller than the number of routing nodes in the network. To our knowledge, this work is the first to study the impact of variant assignment to routing nodes on overall network resiliency.

We present a novel problem, the Diversity Assignment Problem (DAP), which specifies how to optimize overall network resiliency when placing diverse variants that are compromised independently at routing nodes. While DAP is NP-Hard, we show that it is feasible to solve it optimally on a variety of medium-size random network graphs. We also show an efficient algorithm that approximates DAP well for larger

graphs, incurring a relatively small resiliency cost compared with the optimal solution.

To check the applicability of our approach in a real-world setting, we obtained a network graph representative of the global overlay topology used by the above cloud service. Even though this topology was constructed with high availability as the goal (rather than intrusion-tolerance), the optimal variant assignment solution to the DAP ensures a system resiliency that is significantly higher than the resiliency achieved by any of the individual variants.

In real-life settings, routing nodes may be added from time to time to meet increasing system demands. Calculating an optimal solution for the extended network is certainly feasible. However, that solution is likely to require variant re-assignment for many of the existing routing nodes, which may not be feasible in a 24/7 service as downtime for re-configuring nodes is unacceptable. We present an online version of DAP that finds the optimal incremental assignment. When applied to the mentioned global topology, we discover that an important trade-off exists between the resiliency the system achieves and how often the network changes.

We initially choose an application agnostic metric for network resiliency that captures the expected client-to-client connectivity between all pairs. We investigate the advantages of considering the specific resiliency needs defined by the nature of a distributed application running at the clients. Specifically, we show how to find the optimal assignment for the underlying network supporting either the Paxos [6] or Byzantine Fault-Tolerant (BFT) [7] protocols. When applied to the mentioned global topology, we found that an assignment that is tailored to these application requirements can provide higher resiliency than an assignment that focuses on general network resiliency obtained by maximizing the expected client-to-client connectivity. Furthermore, we show how to optimize for a weighted resiliency metric for resiliency metric. Such an optimization offers the ability to designate specific client pairs with higher importance such that the assignment prioritizes these client pairs by offering them higher resiliency.

Our assignments are based on assumptions of accurate information about compromise probabilities of variants. Having inaccurate information results in a different assignment which can impact the resiliency of the system and the confidence of the network operator in the resulted assignment. We analyze and measure in a realistic scenario these two types of errors resulting from inaccurate information to understand the impact of inaccuracies in compromise probabilities on assignment resilience and network operator confidence. Our results show that small inaccuracies in information only result in minor errors in assignment and confidence.

The contributions of this paper are as follows:

- We introduce the Diversity Assignment Problem (DAP). DAP describes how to assign diversity to routing nodes in order to maximize the probability of each client pair being connected.
- We formulate the DAP using mixed integer programming (MIP) [8] and find the optimal solution on random graphs constructed in a manner reminiscent of real overlay topologies. To support larger graphs, we extend this formulation to a fast greedy approximation and demonstrate results that are relatively close to the optimal solution in such larger graphs.
- We extend our approach to optimize network resiliency for a given application's demands, rather than for overall expected client-to-client connectivity, to maximize system resiliency. Additionally, we offer an extension that prioritizes optimization for specific client-to-client connections such that the assignment offers higher resiliency to certain connections while sacrificing resiliency of other connections.
- We analyze the impact of diversity on a real cloud overlay topology and extend our approach to support adding routing nodes to the graph in an online manner to address increased client demand.
- We analyze the loss in resiliency when optimally assigning variants based on inaccurate information about compromises.

The rest of the paper is organized as follows. Section II describes our network and attacker models. Section III presents the general DAP along with an optimal solution. Section IV describes and evaluates a greedy approximation algorithm to solve DAP in larger topologies. Section V shows how resiliency is affected in dynamic topology scenarios. Section VI shows the increased advantage of performing diversity assignment with client application knowledge. Section VII shows how to convert DAP to prioritize specific client pairs. Section VIII analyzes how inaccurate compromise information affects assignment. Section IX lists work related to ours. Section X concludes this work.

II. Model

We describe the model of the network and attacker which we consider in this work. These models are quite general as our approaches can be applied in various networking contexts with various of diversity techniques. Our motivation started with a scenario of cloud services being provided over a global network of datacenters while diversifying operating systems for improved resilience, but we noticed that the core problem is general to any network.

A. Network model

We assume a network topology of routing *nodes* that provide communication to *clients*. We assume no control over the structure of the network topology as this is fixed based on the constraints of the networking context. In an overlay routing context, network links impose overhead to continuously monitor their latency and loss characteristics, thus the degree at each node must be limited while ensuring the entire network is still well connected. Alternatively, in a wireless context, network links are limited by the physical broadcast range of each node. We assume that we have a set of diverse variants and we can configure each routing node with a single variant. Our network goals are to maximize the number of client connections or an application-specific communication requirement of the clients.

B. Attacker model

We assume that there is no way to configure a routing node that meets our network needs while being completely

invulnerable to attacker attempts of compromise. Thus, we adopt a probabilistic attacker model where each variant is compromised with some probability. We capture the benefit of diversity by assuming any pair of variants are compromised independently. We assign a probability that an attacker is able to both find a vulnerability and create a successful exploit against a variant within a given time period, and then any routing node in the network with this variant will become compromised. As our probabilities are with respect to a certain time frame, a full long-term system would need mechanisms to detect and recover compromised variants. We consider such mechanisms as outside the scope of this current work. Our probabilistic model of compromise offers a useful way to reason about an attacker's capabilities and measure a network's resilience. Even in realistic scenarios where an attacker is not modeled well probabilistically, we are still raising the bar for the attacker to ensure the attacker must find vulnerabilities and create exploits for different variants of routing nodes.

We do assume a byzantine tolerant routing protocol is used for routing to ensure that communication can occur between two clients as long as a honest path of routing nodes exists.

III. Diversity Assignment

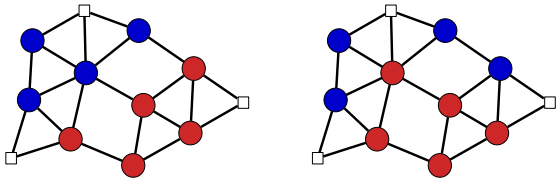
In this section we present the Diversity Assignment Problem (DAP). DAP describes how to assign diversity to routing nodes in order to maximize the probability of each client pair being connected. We then describe existing Mixed Integer Programming (MIP) techniques and how these can be used to solve DAP. Lastly, we show the effectiveness of this technique on a realistic case study topology when compared with randomly assigning diversity.

A. Diversity Assignment Problem (DAP)

We consider a network consisting of a set of nodes N and a set of clients M. A set of connections are defined among these nodes, so we can represent a network as a graph such as the one in Figure 1. Each routing node is assigned a variant from the set of variants V, so there are $|V|^{|N|}$ possible assignments. We denote an assignment of one variant for each node as A. Note that $|V| < |N|$. Each variant $v_k \in V$ is associated with a compromise event e_k in the set of all compromise events E, so $|E| = |V|$. The probability of e_k occurring is $P(e_k)$. These events of compromise are independent,* so for any two compromise events $e_{k'}$ and $e_{k''}$ the following holds $P(e_{k'} \cap e_{k''}) = P(e_{k'}) * P(e_{k''})$.

We measure the goodness of an assignment of variants with the metric *expected client connectivity*. This metric is the expected value of the proportion of client pairs that are connected. To compute this value we consider the set of all possible combinations of compromise events C where $|C| = 2^{|E|}$ (C is the powerset [9] of E). An element $c \in C$ is a subset of the compromise events, E, and corresponds to those compromise events occurring while any other compromise events do not occur. We can compute the proportion of clients connected given that those variants are compromised.

*We make an assumption of independence among compromise events in this work as this simplifies the presentation of the fundamental ideas in this work. We provide analysis of what occurs when compromise events are not highly positively correlated in Section VIII.



(a) (b)

Fig. 1. Example of two assignments on the same topology where routing nodes are circles and clients are squares. We show two possibilities for diversity assignment to nodes where the two variants are red which has a 0.1 probability of being compromised and blue which has a 0.15 probability of compromise. (a) Diversity assignment with 0.838 expected client connectivity. Notice that only one client pair is connected if either red or blue is compromised. (b) Superior diversity assignment that has 0.957 expected client connectivity. Notice that three client pairs are connected if blue is compromised and two client pairs are connected if red is compromised.

We consider two clients to be connected if a path of non-compromised nodes exists between them.

Our goal is to maximize the expected client connectivity of a graph by strategically assigning variants. We call this problem the Diversity Assignment Problem.

Definition 1: The Diversity Assignment Problem is to find the assignment of variants to nodes which maximizes the expected client connectivity. First, for a given assignment A and set of compromised variant events $c \in C$, we define a connectivity function $f_{A,c}(a,b)$ between two clients a and b as:

$$f_{A,c}(a,b) = \begin{cases} \binom{|M|}{2}^{-1} & \text{if clients } a \text{ and } b \text{ are connected by a set of non-compromised nodes} \\ 0 & \text{otherwise} \end{cases}$$

Then, the expected client connectivity is:

$$E\left[\sum_{\{a,b\in M:a<b\}} f_{A,c}(a,b)\right] = \sum_{c\in C}\left(\prod_{e_k\in c} P(e_k)\prod_{e_k\notin c}(1-P(e_k)) * \sum_{\{a,b\in M:a<b\}} f_{A,c}(a,b)\right)$$

The Diversity Assignment Problem is:

$$\text{argmax}_A\left(E\left[\sum_{\{a,b\in M:a<b\}} f_{A,c}(a,b)\right]\right)$$

Theorem 1: The Diversity Assignment Problem is NP-Hard with two or more variants.

Proof: The proof is in Appendix I. ∎

We illustrate the meaning of DAP in Figure 1 with an example topology graph. Figures 1(a) & 1(b) show two ways to assign variants in this graph. Figure 1(b) is the superior assignment as more client pairs are connected given that a single variant is compromised. The superiority of this assignment is also reflected by the expected client connectivity values.

B. MIP approach to DAP

Despite DAP being NP-Hard, many real-world network topologies are of limited size, so finding the optimal solution is of practical interest. To find the optimal solution, we chose to formulate the problem as a MIP and utilize an existing commercial solver, CPLEX [10]. A MIP is a linear program with the addition of integer constraints. The important implication of these integer constraints is that a MIP is not solvable in polynomial time (while a linear program can be), but these integer constraints allow for formulations of many difficult combinatorial problems. Problems from other domains have also resorted to MIP to find optimal solutions to practical problems in the area of operations research [11], [12], [13]. MIP formulations are good for problems where the optimal is desired and no efficient algorithm is known as many MIP solvers [10], [14], [15] employ a variety of techniques to avoid exhaustively searching the entire space of feasible solutions.

Our MIP formulation is seemingly more complex than the mathematical formulation in Definition 1 mainly due to the expression of the function $f_{A,c}(a,b)$ as a MIP. This function's output depends on whether two clients are connected given an assignment and set of compromise events. In the MIP formulation we capture the same connectivity by setting up flow variables on each edge. When considering a specific source client, we count the number of other clients that are connected to this source client with the following constraints on these flow variables. The source client has no incoming flow and unbounded outgoing flow, each other client accepts at most one unit of incoming flow and has no outgoing flow, and each non-compromised node has equivalent incoming and outgoing flow. Compromised nodes have no incoming or outgoing flow, and a node is compromised when the node's variant assignment is included in the set of compromised events being considered. With these flow variables, $\sum_{\{a,b\in M:a<b\}} f_{A,c}(a,b)$ is equivalent to $\frac{1}{2} * \binom{|M|}{2}^{-1}$ multiplied by the total outgoing flow of the given clients for $|M|$ copies of the same graph and flow variables where each graph considers a different source client. Then, we must copy these variables again, once for each possible set of compromise events.

Table I describes each symbol that we use in our MIP formulation. We present the objective function (Equation 19) followed by each constraint (Equations 2-10).

DAP objective:

$$\max_{s,f} \quad \frac{1}{2} * \binom{|M|}{2}^{-1} * \sum_{c\in C, a\in M, x\in N} \left(\prod_{e_i \in c} P(e_i) \prod_{e_i \notin c} 1 - P(e_i) \right) f_{c,a,a,x} \tag{1}$$

We maximize the expected client connectivity of the graph, over all compromise events. The first term ($\frac{1}{2} * \binom{|M|}{2}$) ensures that the result will be out of 1, rather than out of the number of possible connections between clients. The two products ensure that each possible compromise event is weighted by the probability that it happens. The f term is a measure of how much flow the given client a can push out onto the network (specifically, $f_{c,a,i,j}$ measures the amount of flow that started at source client a that travels on edge $\{i,j\}$ in compromise case c). Because of all the constraints below, this is exactly a measure of how many other clients a can connect to.

TABLE I
NOTATION

Symbol	Description
N	Set of routing nodes. As our notation, these are x, y, z, etc. Depicted by circles in figures.
M	Set of client nodes. As our notation, these are a, b, etc. Depicted by squares in figures.
V	Set of variants. Depicted by colors of circles in figures.
E	Set of all compromise events. We index elements of E and V by k as their elements are related such that each e_k corresponds to the compromise event of the variant v_k.
C	Set of all possible compromise event sets, so $\|C\| = 2^{\|E\|}$. Each element $c \in C$ is a set of compromise events ($e \in E$) that are compromised.
$w_{i,j}$	Constants designating that edge $\{i,j\}$ exists. i and j can be either routing nodes or client nodes. Note that clients should not connect directly to other clients, so $i,j \in M \Rightarrow w_{i,j} = 0$. Depicted by lines between nodes in figures.
$f_{c,a,i,j}$	Measures the amount of flow that starts at client node a and travels on edge $\{i,j\}$ in compromise event set c. i and j can be either routing nodes or client nodes. Also, $c \in C$. This must be a non-negative value.
$s_{v,x}$	The variant assignment of routing node x. $s_{v,x}$ is 1 if x is variant v and 0 otherwise.

Variant constraints (I):

$$s_{v_i,x} = \{0,1\}, \quad v_i \in V, \quad x \in N \tag{2}$$

Routing nodes must be either entirely of a variant or entirely not of that variant. Fractional assignments are not allowed.

Variant constraints (II):

$$\sum_{v_i \in V} s_{v_i,x} = 1, \ x \in N \tag{3}$$

Routing nodes must be exactly one variant.

Node flow constraints:

$$\sum_{i\in N\cup(M-\{a\})} f_{c,a,x,i} - \sum_{i\in N\cup\{a\}} f_{c,a,i,x} = 0, c \in C, \ a \in M, \ x \in N \tag{4}$$

The flow (originating at source client node a) entering routing node x must equal the flow (originating at source client node a) exiting routing node x. This is enforced for each of the $|M|$ clients and for each of the $|N|$ nodes, separately. In other words, flow cannot get stuck in the middle of the network; it has to end at client nodes.

Client flow constraints (I):

$$\sum_{x\in N} f_{c,a,x,b} \leq 1, \ c \in C, \ a,b \in M, \ a \neq b \tag{5}$$

A client cannot accept more than one unit of flow from another client. This is so that we can count the total flow out of the source client to get the number of connected clients. Despite this constraint being ≤ 1, it can only take a value of 0 or 1 due to the other constraints and the objective. For the CPLEX solver [10], it is more efficient to enforce fewer integer constraints whenever possible.

Client flow constraints (II):

$$f_{c,a,x,a} = 0,\ c \in C,\ a \in M,\ x \in N \tag{6}$$

Traffic cannot start and end at the same client. In other words, a client cannot send to itself. Note that $\{x, a\}$ is any incoming edge into a.

Client flow constraints (III):

$$f_{c,a,b,x} = 0,\ c \in C,\ a, b \in M,\ x \in N,\ a \neq b \tag{7}$$

A destination client cannot send out flow. So, flow cannot use a client to reach other clients.

Topology constraints:

$$f_{c,a,i,j} \leq (|M| - 1) * w_{i,j},\ c \in C,\ a \in M,\ i, j \in (N \cup M) \tag{8}$$

Any pair of nodes with no edge between them (i.e., $w_{i,j} = 0$) cannot have any flow directly between them. It also underlines the fact that up to $|M|-1$ units of flow originating at the same client can share the same edge.

Variant flow constraints (I):

$$\begin{aligned} &f_{c,a,x,i} \leq (|M| - 1) * \min_{e_i \in C}(1 - s_{v_i,x}), \\ &c \in C,\ a \in M,\ x \in N,\ i \in N \cup M \end{aligned} \tag{9}$$

The amount of flow out of a routing node must be 0 if that node is compromised. It also underlines the fact that no edge can carry more than $|M| - 1$ units of flow from any source client node a.

Variant flow constraints (II):

$$\begin{aligned} &f_{c,a,i,x} \leq (|M| - 1) * \min_{e_i \in c}(1 - s_{v_i,x}), \\ &c \in C,\ a \in M,\ i \in N \cup M,\ x \in N \end{aligned} \tag{10}$$

The amount of flow into a node must be 0 if that node is compromised. It also underlines the fact that no edge can carry more than $|M| - 1$ units of flow from any source client node a.

C. DAP on the case study topology

We investigate the benefit of optimal diversity assignment on a realistic overlay network topology. Then, various assignments of diversity are shown on the case study topology with their corresponding expected client connectivity. We show assignments for DAP with increasing number of variants being used, and we investigate random assignments as a comparison with the optimal solution.

For a case study topology, we took a connectivity graph from a cloud network provider [5]. The nodes of the graph represent data centers located around the globe. Each node is assigned a single variant which means that the overlay routing element at that data center will utilize the selected variant. The edges of the graph represent overlay connectivity used on that cloud to connect the different data centers. This connectivity is provided by a number of Internet Service Providers at each data center. The clients in the graph represent either clients external to the cloud or infrastructure components of the cloud. Each client has multiple connections to the cloud to avoid a single point of failure. In this example, we use three connections as that level of connectivity was quite prevalent in that network. This connectivity graph was designed with resiliency in mind, and without any consideration for diversity.

We assume some hypothetical scenario with three diverse variants represented by red, blue, and green having a 0.1, 0.15, and 0.2 probability of being compromised over some arbitrary period of time, respectively. Note that this example, while simplistic, provides an interesting insight into the benefits and risks of diversity. †

Figure 2(a) shows the optimal solution when only a single variant can be used. All the nodes are assigned with the least vulnerable variant. This corresponds to the situation where no diversity is used. The resulting network achieves an expected client connectivity of 0.9.

Figure 2(b) shows the optimal solution when two variants can be used. Each node is assigned with either of the two least vulnerable variants. The resulting network achieves an expected client connectivity of 0.985. Note that this is better than either variant by itself.

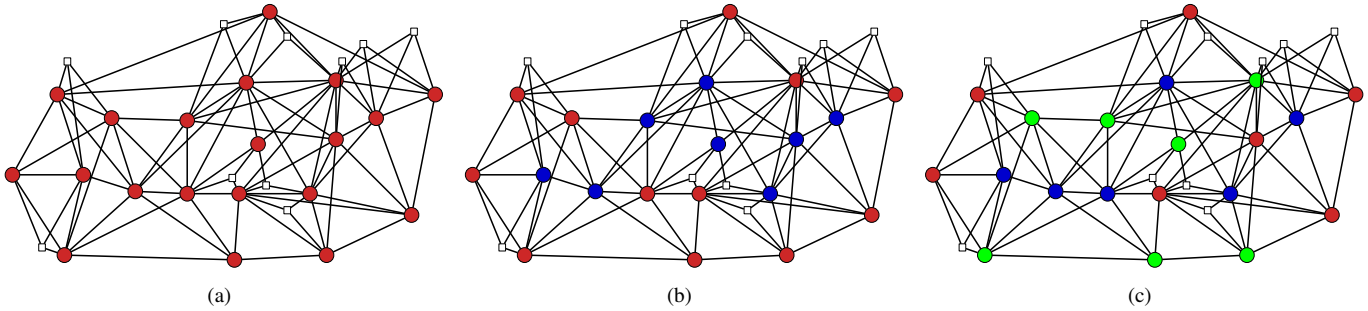
Figure 2(c) shows the optimal solution when three variants can be used. The resulting network achieves an expected client connectivity of 0.997. Notice that the optimal solution finds an assignment where any single variant is capable of connecting all clients. By adding a third, more vulnerable variant actually makes the system significantly more resilient.

As stated before, in this example, each client is connected to three routing nodes. If clients do not have at least three potential entry points into the network, then the availability of the connection is limited by the variants of the routing nodes that they are connected to. For example, if each client only connects to a single routing node, that connection would fail if either of the entry-point routing nodes is compromised. This is much more likely to occur than if there are three such entry-point routing nodes for each client, requiring at least three routing nodes to be compromised to cut the connection.

In this example, including variants *that have a higher but independent probability of being compromised* improves the overall system resiliency. This may be counterintuitive, as adding weaker components to a system usually makes it weaker, not stronger. The independence of the different variants and the overall robustness of the network mean that adding additional, more vulnerable variants makes a system more resilient.

As discussed earlier, random assignment could be used instead of the optimal MIP approach. One might expect this approach to do well, since randomness often helps in adding diversity to systems. However, this does not necessarily lead to a good result. An example graph can be seen in Figure 3(b). This graph achieves an expected client connectivity of only 0.811, much worse than any of the other three graphs. In fact, it barely outperforms the worst of the three variants. This example graph comes from the bottom 1% of possible assignments and is given as an example of what could occur if the diversity assignment is not considered carefully.

†The purpose of these values is to give preference to one variant over another and to quantify an estimate of the system resiliency with diversity. While we select numbers to illustrate the main concepts, the resulting assignment would not be significantly different if other values were selected.



(a) (b) (c)

Fig. 2. Optimal assignments on case study topology: (a) one variant assignment achieves 0.9 expected client connectivity, (b) two variants assignment achieves 0.985 expected client connectivity, (c) three variants assignment achieves 0.997 expected client connectivity.

Figure 3(a) is a histogram created with data from 100,000 random assignments on the case study topology. For this data set, the minimum and maximum are 0.751 and 0.988 respectively. The mean is 0.931 and the median is 0.937. As can be seen, most of the random assignments perform better than if the best variant is used by itself $(0.937 > 0.9)$. However, very few of the random assignments come close to performing as well as the optimal assignment found by MIP.

The optimal solution of 0.997 expected client connectivity exists while the best random solution out of the 100,000 random assignment shown in Figure 3(a) was 0.988 expected client connectivity. Thus, even the best random solution out of numerous trials does not achieve the optimal solution. We define *expected client disconnectivity* to be the expected probability that communication between a client pair is broken, and this value is equivalent to (expected client disconnectivity) = 1 - (expected client connectivity). In terms of expected client disconnectivity the best random solution is 0.012 while the optimal solution is 0.003, so a client-to-client connection is broken four times less often with the optimal assignment.

Interestingly, the difference between what the optimal solution provides and the probability that at least one of the variants is non-compromised provides a metric for the quality of the connectivity resiliency of the graph.[‡] Ideally, we would want this distance to be zero, as in Figure 2(b) and Figure 2(c) of the provided example.

D. Near-optimal assignments on case study topology

Here, we aim to further understand why the problem is difficult since the optimal solution is several factors better than random assignments from the perspective of the expected client disconnectivity. To achieve this, we compute the set of all assignments near the optimal solution in terms of expected client disconnectivity. The number of assignments found compared to the size of the search space further supports our claim that random assignments are typically much worse than the optimal assignment. Thus, applying techniques of this work to search for optimal assignments is important for any network aiming to achieve high resilience through diversity.

We search for solutions within a *disconnectivity factor* of the optimal solution. This value is computed from a given expected client disconnectivity as follows (disconnectivity factor) = (expected client disconnectivity) / (OPT) where OPT is the optimal expected client disconnectivity. Intuitively, a disconnectivity factor of two for an assignment implies that clients on average are disconnected twice as much as the optimal assignment.

[‡]Thanks to Bob Balzer for this observation.

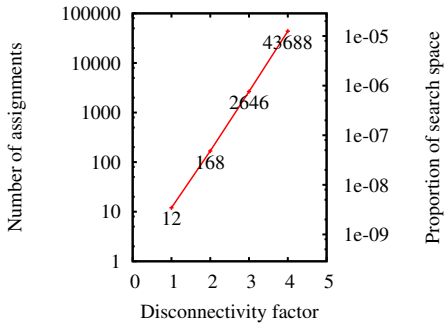


Fig. 4. Number of solutions within a given disconnectivity factor bound for the case study topology.

Exhaustive search of the entire search space is prohibitively expensive for the three variant case, and we could not use this strategy to find all near-optimal solutions. However, we were able to find all solutions within a factor of optimal by leveraging advanced features of MIP solvers. After finding an optimal solution, the solver can be set to continue searching for solutions. The solver avoids exhaustively searching the entire space by eliminating large portions of the search space through its branch and bound techniques. Given that the number of solutions found is small, this procedure is quite efficient.

Figure 4 shows the number of solutions within a small factor of the optimal solution for the three variant scenario (note the log-scale of the y-axis). We show the proportion of the search space that these solutions represent on the right y-axis. The proportion of the search space indicates the probability that a random assignment has of achieving an assignment within a small factor of the optimal solution. Thus, a random assignment has a probability of $3 * 10^{-9}$ to achieve optimal, so that would require on the order of a billion topologies to be assigned and evaluated to find an optimal solution. The visually linear trend in this figure implies an exponential trend in the data due to the logscale of the y-axis. Thus, the number of solutions within a factor of optimal decreases exponentially with respected to decreasing factor, and this implies searching exponentially more assignments to expect to find such a solution.

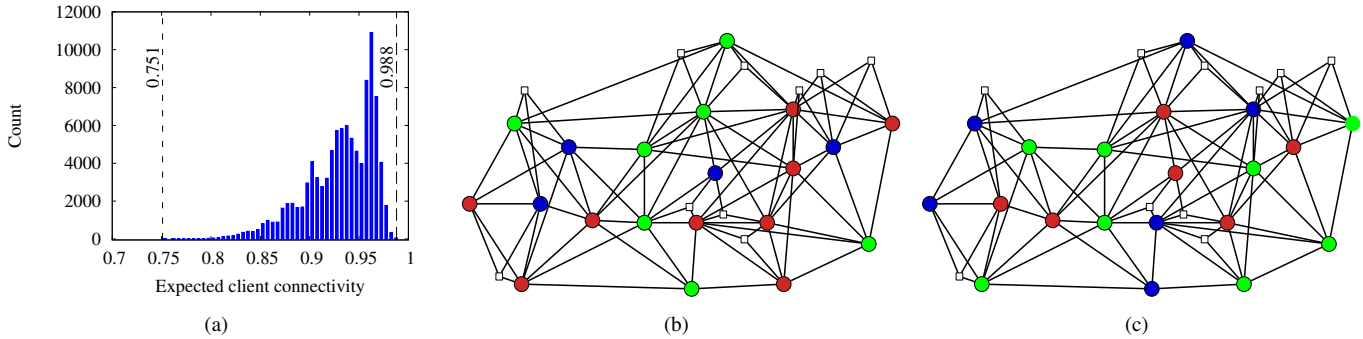


Fig. 3. Random and greedy assignments on case study topology: (a) histogram of expected client connectivity of 100,000 random assignments with vertical lines displaying the lower and upper bounds, (b) random assignment achieving 0.881 expected client connectivity, and (c) greedy assignment achieving 0.992 expected client connectivity.

IV. Scaling Diversity Assignment

DAP is not tractable for large topologies since DAP is NP-Hard (see Theorem 1). To scale to larger topologies, we sacrifice optimality in order to ensure the algorithm completes within a polynomially-bounded time. In this section we present the Approximate DAP (A-DAP), a greedy approach to A-DAP, an example on the case study topology, and an evaluation on random topologies.

A. Approximate DAP (A-DAP)

A-DAP is similar to DAP, but A-DAP does not require that the problem be solved optimally. By relaxing this condition, we aim to find algorithms that run in polynomial time which are able to find large values of expected client connectivity. We do not formally define any restrictions on the goodness of the approximations as it is an open problem of whether a reasonable bound can be placed on the expected client connectivity achieved by a deterministic polynomial time algorithm. Instead, we used random topologies to validate the goodness of expected client connectivities achieved by a greedy approach to A-DAP when compared with the optimal.

B. Greedy approach to A-DAP

Our greedy approach incrementally assigns nodes to variants. At each incremental assignment, the algorithm considers several candidate assignments and selects the one which provides the best immediate results. For a candidate set of incremental assignments we consider sets of nodes which can connect a client pair by a variant, so we consider at most $\binom{|M|}{2} * |V|$ candidate variant assignments. For a given client pair a and b and variant i, we compute the minimal number of unassigned nodes which must be assigned i to connect a and b by nodes assigned i. After this computation we have two values: the increase in expected client connectivity α and the number of newly assigned nodes β.

Given a set of candidate assignments that each have an α and β value, we select the one which maximizes $\frac{\alpha}{\beta}$. It is obvious why we want to find large α values, but it is equally important to ensure the β value is small as well. Smaller values of β allow for more nodes to remain unassigned and to be used to connect more client pairs by other variants in future assignments. This approach is analogous to the greedy choice in bin packing, as we select items with the highest payoff versus weight ratio to ensure that items are selected that increase overall payoff while allowing for more items to be picked in the future. Note, that $\beta = 0$ is a trivial case where the candidate is simply removed from consideration as the client pair is already connected via the considered variant.

The pseudo-code of the algorithm is shown in Algorithm 1. Each iteration of the while loop (Line 3-19) creates a set of candidate variant assignments (Line 7), then selects the best candidate (Line 11-15), and lastly applies the assignment of that candidate to the topology (Line 18). This algorithm completes when no further client pairs can be connected by a variant, and the algorithm is guaranteed to complete in a bounded number of iterations since each step connects at least one new client pair via a variant (at most $|C|*|M|^2$ iterations).

C. A-DAP on the case study topology

We consider the same scenario as in Section III-C with three variants. Figure 3(c) shows the assignment found by our greedy solution which achieves 0.992 expected client connectivity. Notice that all clients are connected via just the blue or yellow variants. However, two clients remain disconnected from the rest if only the red variant is uncompromised. The optimal solution found with the MIP formulation finds an assignment which connects all clients as long as any single variant is uncompromised. This loss of expected client connectivity is due to the greedy algorithm making choices in the early steps of the algorithm to connect clients via blue and yellow variants (the more resilient variants) which leaves fewer choices to connect clients via the red variants. The greedy approach for the A-DAP took 0.38 seconds to complete while the MIP approach for the DAP took 396.13 seconds to complete. With far less computational requirements, the greedy algorithm does outperform the best of the 100,000 random assignments (0.988 client connectivity) and comes close to the optimal solution.

D. A-DAP on random topologies

We present a methodology followed by results to answer the following questions of interest about the performance of the greedy algorithm for the A-DAP:

1) How does the goodness of the assignment of the greedy algorithm compare to other algorithms (random assignment and optimal) for the DAP on typical topologies?
2) How does the running time of the greedy algorithm for the A-DAP and the MIP approach for the DAP vary with typical topologies created with different parameters?

Algorithm 1 Greedy assignment algorithm

Variables

CPVC: Client Pair and Variant Combinations
VA: Variant Assignment
DVA: Delta Variant Assignment
CG: Connectivity Gain
BCG: Best Connectivity Gain
DVA: Delta Variant Assignment
BDVA: Best Delta Variant Assignment
α: Tunable parameter which affects the trade-off between increasing connectivity and minimizing the size of the DVA set

Functions

$f(\cdot,\cdot)$: Minimal set of unassigned overlay nodes that must be assigned a particular variant to connect a particular client pair
$g(\cdot)$: Overall connectivity for a particular variant assignment

Algorithm

CPVC := $M \times M \times V$
VA := $\emptyset$
while CPVC $\neq \emptyset$ **do**
  BCG := 0
  BDVA := $\emptyset$
  for all $x \in$ CPVC **do**
    DVA := $f(x, \text{VA})$
    if DVA $= \emptyset$ **then**
      CPVC := CPVC $- x$
    else
      CG := $\frac{g(\text{VA} \cup \text{DVA}) - g(\text{VA})}{|\text{DVA}|^{\alpha}}$
      if CG $>$ BCG **then**
        BDVA := DVA
        BCG := CG
      end if
    end if
  end for
  VA := VA $\cup$ DVA
end while

3) What are trends in the expected client connectivity over all the assignment algorithms when varying topology parameters?

We use expected client connectivity and running time to evaluate each algorithm. Expected client connectivity is a measure of how well the algorithm performs. Running time is a measure of how quickly the algorithm will terminate with an expected client connectivity.

We generate random topologies in a similar way to random wireless topologies. That is, we place the desired number of nodes and clients randomly inside a two-dimensional box. Then, based on a density parameter, we give each node and client a range. All nodes and client within the range have an edge between them. The density parameter is the average number of nodes each node or client is connected to. Note that client to client edges are not added. We can create many random topologies given a number of nodes and a density value. We chose to limit the number of nodes in order to ensure that the optimal value could be calculated for comparison. Topologies constructed in this way are obviously representative of wireless contexts, but they are also quite similar to overlay topologies, because overlay topologies include many short, well-behaved links.

Given topology parameters, we create 30 random topologies and run the three algorithms on these topologies. We average the expected client connectivity and running times obtained for each algorithm over the 30 runs. For the running time values of the MIP formulation, it is important to note that we use the software package CPLEX with a quad-core 3.4 Ghz Intel processor which does leverage all cores.

The results are shown in Figure 5. We describe how they answer each of the initial questions that we proposed.

Question 1. The goodness of an algorithm's assignment is the expected client connectivity. This is upper-bounded by the optimal value (which the MIP approach always achieves). The greedy algorithm outperformed the random assignment and was quite close to the optimal value, independent of varying either density (Figure 5(a)) or the number of nodes (Figure 5(c)).

Question 2. The running time of the greedy algorithm is on the order of milliseconds, which is barely visible when compared to the running time of the MIP-based approach. Figure 5(b) shows the MIP approach running time for varying density values. The running time is low for small density values since most variant assignments result in poor expected client connectivity, allowing the branch-and-bound algorithm of CPLEX to avoid searching the majority of variant assignments. The running time is also low for high density values since a dense graph has many possible optimal assignments and the branch-and-bound algorithm can terminate early after finding any of them. Thus, the problem is the most difficult for networks with moderate density values. The running time of both algorithms when varying the size of the network is shown in Figure 5(d). The MIP approach running time grows nearly linearly over these input parameters, but this relationship is potentially exponential according to Theorem 1. The MIP approach running time is still significantly greater than the greedy approach.

Question 3. The trend of expected client connectivity is similar among all three algorithms. The expected client connectivity increases as density increases (Figure 5(a)), which is expected since more edges allow more possibilities for clients to become connected. The expected client connectivity decreases as the number of nodes increases (Figure 5(c)). By keeping the density constant and increasing the number of nodes, the graph becomes less connected and therefore less resilient.

From these results we see that the greedy algorithm outperforms the random algorithm while being quite close to the optimal solution, and the greedy algorithm is far more efficient in terms of running time and is polynomially-bounded while the MIP formulation is not. Hence, on larger topologies where the MIP formulation cannot be computed, the greedy algorithm is a decent substitute. Another interesting result is that the expected client connectivity decreases with more

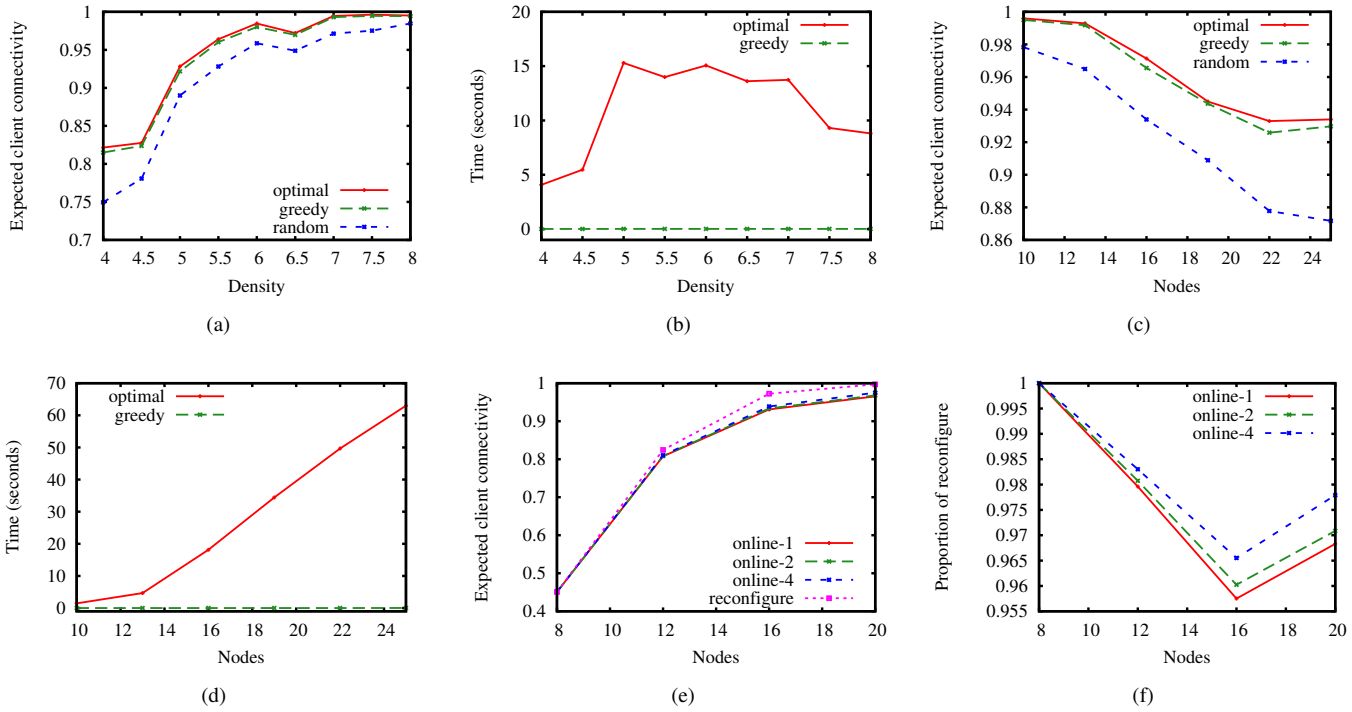


Fig. 5. Experiments for both (a,b,c,d) comparing random, optimal, and greedy algorithms and (e,f) comparing online assignments on dynamic topologies. Figures (a) and (b) show results of random, optimal, and greedy algorithms on random topologies with 25 nodes and varied density. Figures (c) and (d) show the random, optimal, and greedy algorithms on random topologies with 6 density and varied nodes. Figure (e) shows expected client connectivity achieved by reconfigure and online versions as the network grows while Figure (f) shows the expected client connectivity achieved by each online version, as a proportion of the optimal value achieved by reconfiguring (1.0 on the graph).

nodes when keeping the density constant. So, the density or node degree must increase to retain high levels of expected client connectivity when the number of nodes increases in the topology.

V. Diversity Assignment for Dynamic Topologies

In practice, networks typically do not remain static throughout their lifetime. Instead, an initial setup is deployed and over time, nodes are dynamically added. One trivial way to leverage diversity in an online scenario is to solve DAP every time a change in the topology occurs. However, for many classes of diversity it is highly expensive or even prohibitive to reassign an existing node of one variant to a different variant as it may be difficult to revoke access from an administrator or expensive to reinstall a new diverse software. A more realistic solution is to always keep the existing variant assignment and just assign variants to the newly added nodes.

We describe the specific model which captures our assumptions. Then, we describe an approach to solve this problem. Lastly, we evaluate this approach for an online scenario.

A. Online DAP (O-DAP)

We assume that there is some variant assignment that exists for a set of nodes which we denote by A'. A new set of nodes are added to the topology with given links to existing nodes in the network. We assume that there is no knowledge of future topology changes, so we cannot anticipate where new nodes may be added, which is an assumption that is realistic in practice. We seek a variant assignment, A, which retains all of the variant assignments of A'. We denote this problem as the Online Diversity Assignment Problem (O-DAP) with formal details in Definition 2.

Definition 2: The Online Diversity Assignment Problem extends DAP by adding additional constraints. There exists some set of nodes which have already been assigned variants, and this existing assignment is denoted by A'. We are using the notation $A' \subset A$ to convey that the assignment A must retain the assignment of A'. The assignment A does have the freedom to assign variants in any way to those new nodes added to the network which are not contained in the assignment A'. Reusing notation from Definition 1, we can define the Online Diversity Assignment Problem as:

$$\begin{aligned} &\text{argmax}_A && \left(E\left[\sum_{\{a,b \in M : a<b\}} f_{A,c}(a,b)\right]\right) \\ &\text{subject to} && A' \subset A \end{aligned}$$

Theorem 2: The Online Diversity Assignment Problem is NP-Hard with two or more variants.

Proof: Let $A' = \emptyset$, and then O-DAP is equivalent to DAP. Theorem 1 states that DAP is NP-Hard. ■

B. MIP approach to O-DAP

We detail a MIP approach for O-DAP as it is typically easy to solve O-DAP optimally because the number of nodes which are added to a network at once is usually small. Given that x nodes are added to the network and x is small, then the search space, $|V|^x$, is reasonably small as well. Exhaustive search by checking all possible variant assignment combinations of the x new nodes could be used. However, as we already have a MIP formulation available to us, it is simple to reformulate

the MIP that optimally solves DAP to optimally solve O-DAP. Specifically, we add the following constraint to the same MIP formulation for DAP from Section III-B.

Online variant constraints:

$$s_{v_i,x} = 1, \quad \langle x, v_i \rangle \in A' \tag{11}$$

Nodes that have been assigned previously by A' (elements in A' are two tuples denoting a node and its corresponding variant assignment) must keep that variant assignment.

Theorem 2 states that O-DAP is NP-Hard. In scenarios where the number of nodes being added dynamically is large, it is possible to extend the greedy approach for A-DAP into an online version that approximates O-DAP.

C. O-DAP on the case study topology

The expected client connectivity of DAP is always greater than or equal to that of O-DAP for the same topology, since O-DAP only adds constraints to DAP. For a real deployment this means that reconfiguring all of the variants to be optimal when each dynamic change occurs always results in equal or better expected client connectivity compared with an online version where existing variants cannot be reconfigured. We measure this degradation in expected client connectivity for this evaluation.

The size of the incremental node additions influences the resulting expected client connectivity. A network which does many additions of just a few nodes per topology change will suffer more in expected client connectivity than a network which adds many nodes per topology change. Networks which add many nodes at once allow O-DAP to consider more combinations of variant assignment choices. We consider the following two scenarios for dynamic topologies in our evaluation:

- **reconfigure**: DAP is solved and that solution is applied to all nodes in the network. As reconfiguring is typically an unreasonable approach in practice, we use this as a baseline for comparison with the online approach.
- **online-x**: Nodes are added to the network x at a time, and O-DAP is solved where the variants of existing nodes in the network cannot be changed.

We evaluate these strategies with the following scenario on our case study topology. We initialize a scenario topology from our case study topology by selecting 8 of the 20 nodes. Next, we solve the DAP for the scenario topology. Then, we add nodes to the scenario topology based on the strategy being used (i.e., x for online-x) until all 20 nodes are in the scenario topology. We keep the order in which nodes are added to the scenario topology consistent across different strategies. For example, the first four nodes added one at a time in online-1 will be the same nodes added all at once in the first iteration of online-4. Note that, while the topologies will match, the variant assignments may differ. We repeat this process for 30 different scenarios, randomly varying which nodes are in the initial topology and the order in which the remaining nodes are added, and show averages over these 30 scenarios.

Figure 5(e) shows the results for the evaluated strategies. From this figure, it is evident that the online strategies achieve less expected client connectivity than the reconfigure strategy. To better compare these strategies we show Figure 5(f), which instead of showing absolute expected client connectivity, shows the proportion of the expected client connectivity achieved by the online versions to the expected client connectivity achieved by the reconfigure strategy, which is optimal. The online strategies always achieve at least 95% of the reconfigure strategy. More dynamic strategies reduce client connectivity, but in our experiment this degradation was never more than 1% when comparing online-1 and online-4. The downward then upward trend (V-shape) of this figure is due to the following: the initial downward trend is due to the online strategies diverging more and more from reconfigure strategy at larger node values, the latter upward trend is due to the general high connectivity in the topology which results in any online assignment strategy being close to the reconfigure's optimal assignment as the network becomes fully assigned.

These results indicate that diversity is not limited to static deployments, but that diversity can also be applied effectively when networks are dynamic. However, a trade-off exists; reconfiguring the entire network is costly but it yields the optimal expected client connectivity. It is up to the system designer to judge the correct balance between resilience and reconfiguration cost. For systems with very high resiliency goals, this reconfiguration may be necessary. When the highest resiliency is not necessary, the O-DAP approach can be utilized to eliminate the costs of reconfiguring nodes while sacrificing resiliency. In our experiments, we observed that the O-DAP approach achieved resiliency no worse than 95% of optimal.

VI. Diversity Assignment for Specific Applications

Certain distributed systems that maintain consistent state pride themselves on their ability to tolerate part of the system failing. State machine replication protocols with this property include Paxos [6], Byzantine Fault Tolerance (BFT) [7], Prime [16], and Aardvark [17], where Prime and Aardvark give additional performance guarantees even while the system is under attack. These protocols explicitly state their assumptions about the proportion of replicas that must be correct for safety and liveness properties to hold. However, an equally important consideration is that a sufficient number of correct replicas must be able to communicate with each other via the underlying network. If we view the state machine replicas as clients of the underlying network, then applying diversity to the network improves the resiliency of the overall system.

We use these state machine replication protocols as an example of how to customize DAP for a specific client application. State machine replication protocols have specific connectivity needs among replicas that must be satisfied to ensure safety and liveness. We show how DAP is customized to better ensure the network meets these requirements, and we show how such customization can be helpful in a realistic scenario. The steps we take here to customize DAP can be followed to create other versions that meet the specific connectivity needs of other distributed systems.

The expected client connectivity from DAP maximizes the expected value of the proportion of client pairs that are connected. This is a reasonable metric for resiliency of many

applications, and it could even work well for state machine replication in certain scenarios. However, an approach that takes into account the connectivity requirements of the specific application (in this case, state machine replication) may result in higher overall resiliency. We refine DAP to exactly match the needs of a replicated state machine protocol by maximizing the probability that a specific sized connected component exists among the replicas.

A. Connected Component DAP (CC-DAP)

The goal of this algorithm is to optimize the probability that g clients can communicate with each other. The connected component size g can be derived from the specific state machine replication protocol. We denote this problem as the Connected Component Diversity Assignment Problem (CC-DAP) with formal details in Definition 3 (we use the notation from Table I). Unsurprisingly, this problem is also NP-Hard as stated in Theorem 3.

Definition 3: The Connected Component Diversity Assignment Problem is to find the assignment of variants to nodes which maximizes the probability of a component of clients being connected. First, we define the random variable X_A which is the size of the largest connected component of clients given a variant assignment A. This variable is random as it depends on the random events E. Then, the Connected Component Diversity Assignment Problem is:

$$\text{argmax}_A \left(P(X_A \geq g)\right)$$

Theorem 3: The Connected Component Diversity Assignment Problem is NP-Hard with two or more variants.

Proof: The proof is in Appendix II. ∎

B. MIP approach to CC-DAP

For the MIP formulation we keep the constraints in Equations 2-10 from Section III-B, reformulate the objective function, and add new constraints. Our new objective and constraints include new variables which are used to keep track of which subset of clients are used for a connected component $\beta_{c,a}$ as well as variables to check if the connected component is large enough α_c. We describe the purpose of the new objective and each new constraint in detail to show how it captures the CC-DAP problem.

CC-DAP objective:

$$\max_{s,f,\alpha,\beta} \quad \sum_{c \in C} \left(\prod_{e_i \in c} P(e_i) \prod_{e_i \notin c} 1 - P(e_i) \right) \alpha_c \tag{12}$$

We maximize the probability that a g-sized connected component exists, over all compromise events. The two products ensure that each possible compromise event is weighted by the probability that it happens. α_c is 1 if a connected component of size g is present under compromise event c and 0 otherwise.

Component constraint (I):

$$\alpha_c = \{0, 1\},\ c \in C \tag{13}$$

A g-sized connected component either exists under compromise event c, or it does not.

Component constraint (II):

$$\beta_{c,a} = \{0, 1\},\ c \in C,\ a \in M \tag{14}$$

$\beta_{c,a}$ is 1 if client a is in the g-sized connected component under compromise event c, and 0 otherwise.

Component constraint (III):

$$g = \sum_{a \in M} \beta_{c,a},\ c \in C \tag{15}$$

A valid connected component under compromise event c must be of size g. In any other case, this constraint will not be met. Note, if a larger connected component could exist, this constraint ensures that only g clients are considered, which is required for other constraints.

Component flow constraint (I):

$$f_{c,a,x,b} \leq \beta_{c,b},\ c \in C,\ a, b \in M,\ x \in N, a \neq b \tag{16}$$

A client b, in the connected component under compromise event c, cannot accept more than one unit of flow from another client a. If b is not in the connected component, it will not accept any flow.

Component flow constraint (II):

$$f_{c,a,a,x} \leq (g-1) * \beta_{c,a},\ c \in C,\ a \in M,\ x \in N \tag{17}$$

A client a, in the connected component under compromise event c, cannot send more than $g-1$ units of flow, enough for every other client in the connected component. If a is not in the connected component, it will not send any flow.

Component satisfaction constraints:

$$g * (g-1) * \alpha_c = \sum_{a \in M, x \in N} f_{c,a,a,x},\ c \in C \tag{18}$$

If there exists a g-sized connected component under compromise event c, then there are a total of $g * (g-1)$ units of flow in the network. If no such connected component exists, the total flow is 0.

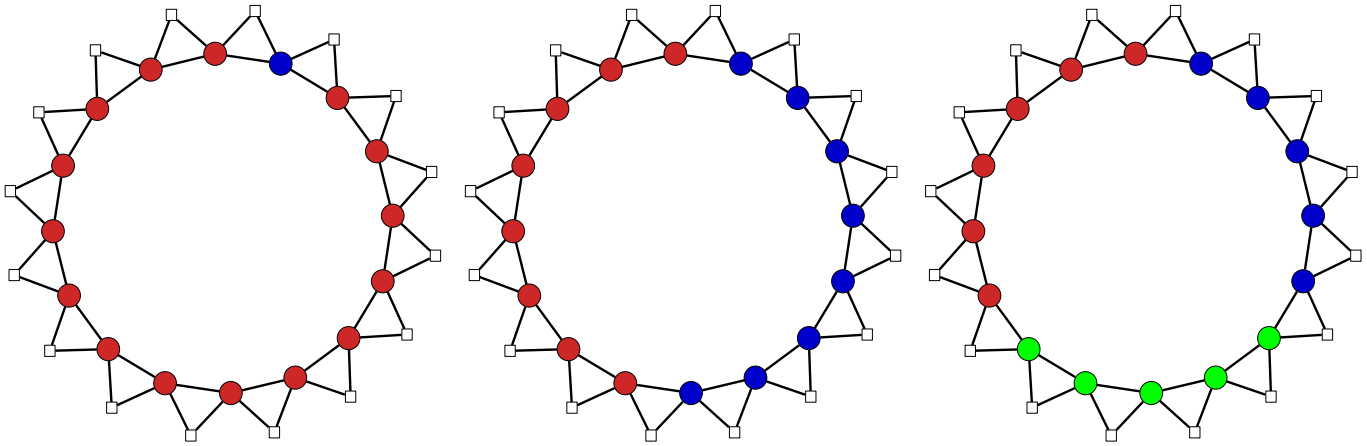
C. CC-DAP on example topology

We provide a quick example on a topology which is contrived to show the advantage of using CC-DAP for applications such as Paxos and BFT. In the following subsections we show this on the case study topology as well.

Figure 6(a) shows the configuration as well as the optimal assignment for DAP. The optimal solution connects all clients by the strongest variant and is able to connect one additional client pair by the second strongest variant.

Figure 6(b) shows with the same configuration the optimal assignment for CC-DAP with a connected component size of 9. For 16 replicas, 9 is the smallest required connected component for Paxos to make progress. The optimal assignment is able to ensure a connected component of 9 with the strongest variant and second strongest variant independently. That is, as long as either the red or blue variant are not compromised Paxos will make progress.

Figure 6(c) shows the optimal assignment for CC-DAP with a connected component size of 11 which is appropriate for BFT with 16 replicas. Here, the optimal assignment ensures 11 clients are connected when any single variant is compromised.



(a) (b) (c)

Fig. 6. Ring topology of 16 clients and 16 routers which highlights difference between optimizing (a) expected client connectivity, (b) probability of Paxos progress, and (c) probability of BFT progress.

TABLE II
VALUES OF THREE METRICS FOR RING TOPOLOGY FOR THREE ASSIGNMENTS THAT EACH MAXIMIZE THEIR OWN METRIC

Assignment	Expect client connectivity	Paxos probability of progress	BFT probability of progress
Figure 6(a)	**0.9004**	0.9	0.9
Figure 6(b)	0.831	**0.985**	0.765
Figure 6(c)	0.787	0.941	**0.941**

That is, if just red, just blue, or just green variants are compromised, then BFT will make progress.

Table II shows the values of each metric for each assignment. Its important to notice how poor certain metrics are when they are not being optimized. Thus, this example shows the value a network can provide when knowing the application being run among the clients.

D. CC-DAP for Paxos

Scenarios where DAP connects all client pairs by every variant individually are trivial for CC-DAP, since an optimal DAP assignment is also an optimal CC-DAP assignment. Thus, we slightly change the setup from Section III-C to ensure a non-trivial comparison between DAP and CC-DAP. In the topology we use for Paxos, we add a new variant v_4 where $P(e_4) = 0.25$ represented in the figures by the color yellow. In the topology we use for BFT, we start with the topology used for Paxos and add new connections between clients and routing nodes. BFT requires this extra modification of including new connections since the nature of BFT requires larger connected components.

Paxos maintains consistent state given that there are at most f_s fail-stop failures when using a total of $n = 2f_s+1$ replicas. In the Paxos scenario, we assume replicas may be partitioned from each other due to attacks on the routing nodes. A client being partitioned from the others is equivalent to a fail-stop failure. For the purposes of this example, we do not consider any other forms of failure, that is, the network may fail but the replicas themselves do not fail. Given that we have 10 replicas in total, this implies that $f_s = 4$. As a result, the required connected component size is $g = n - f_s = 6$.

Figure 7(a) shows the assignment when using the MIP approach for CC-DAP while Figure 2(c) from before shows the assignment when using the MIP approach for DAP. In Figure 7(a), the probability that 6 of the clients will be able to communicate is 0.99925 with an expected client connectivity of 0.9675. In contrast, in Figure 2(c), the probability that 6 of the clients will be able to communicate is only 0.997 while having an expected client connectivity of 0.997 as well. In essence, CC-DAP is able to sacrifice some of the expected client connectivity to increase the probability that a connected component of the desired size will be present.

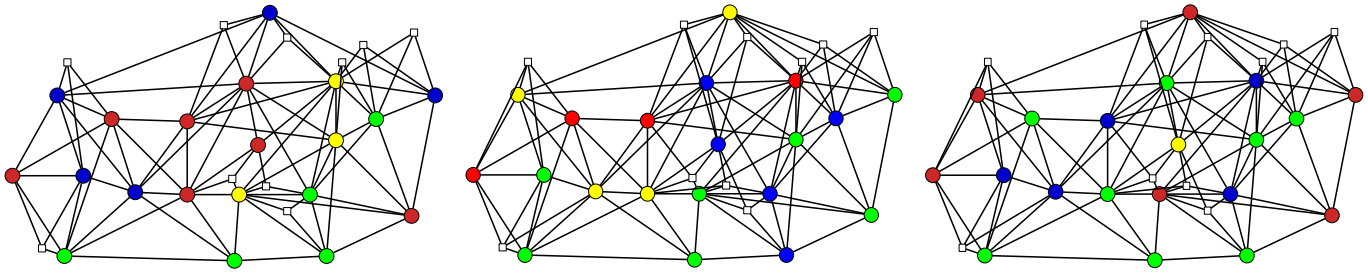
E. CC-DAP for BFT

BFT tolerates up to f Byzantine failures when using a total of $n = 3f + 1$ replicas. We will view these f failures as a combination of f_b, Byzantine replicas, and f_s, fail-stop replicas (indistinguishable from replicas that have been partitioned away). The choice of values for f_b and f_s are left to the system designer. There is trade-off between f_b and f_s, governed by the trustworthiness of the replicas vs. the trustworthiness of the network routing nodes, but further details are beyond the scope of this paper. For our example, we choose $f_b = 1$. Given that we have 10 replicas in total, implying that $f = 3$, the system can tolerate two replicas being partitioned away ($f_s = 2$) and still tolerate one Byzantine fault. As a result, the required connected component size is $g = n - f_s = 8$.

For the results of assignments for BFT, we observe a similar trend to the results of the Paxos scenario. Figure 7(b) shows the assignment when using the MIP approach for CC-DAP that achieves a probability of 0.99925 that 8 of the clients communicate. Figure 7(c) shows the assignment when using the MIP approach for DAP which has only a probability of 0.997 that 8 of the clients communicate.

VII. OPTIMIZING CLIENT TRAFFIC PATTERNS

Specific client connectivity requirements were considered in the previous section. Those connectivity requirements were



(a) (b) (c)

Fig. 7. Assignments illustrating effectiveness of application-specific assignments with CC-DAP: (a) assignment with CC-DAP where Paxos has a probability of 0.99925 to make progress and expected client connectivity is 0.9675, (b) assignment with CC-DAP where BFT has a probability of 0.99925 to make progress and expected client connectivity is 0.9806, and (c) assignment with DAP where BFT has a probability of 0.997 to make progress and expected client connectivity is 0.9975. Note the differences in topologies as (a) has 3 connections from each client to servers while (b) and (c) have 4.

based on ensuring that cliques of communication exist even after router compromises occur. Here, we offer optimization based on client traffic patterns. This offers a selection of importance for certain client pairs based on the amount of traffic or monetary value of the traffic. The choice between these two types of utility is based on the type of network and information available to the network. Instead of finding cliques in CC-DAP or treating all client pairs equally in DAP, we allow arbitrary selection of value for each client pair.

The network can understand the utility gained for offering communication between each pair of clients by observing the quantity of traffic between client pairs or based on payment the network may receive for delivering data between client-pairs. By performing diversity assignment based on this information, the network can better maximize utility by ensuring more important client pairs have communication with higher tolerance to router compromises.

A. Weighted DAP (W-DAP)

This problem has a similar definition to DAP with the exception that traffic weights are included. Each client pair has a given traffic weight which allows us to compute a weighted expected client connectivity for a given assignment. W-DAP is to maximize this weighted expected client connectivity instead of expected client connectivity. We formulate this problem in Definition 4.

Definition 4: The Weighted Diversity Assignment Problem is to find the assignment of variants to nodes which maximizes the weighted expected client connectivity. Let $T_{a,b}$ be the chosen weight value by the network operator for a client pair a, b. Let T^{sum} be the sum of all weight values for each pair of clients $T^{sum} = \sum_{\{a,b \in M : a<b\}} T_{a,b}$, and this is used for normalizing weighted expected client connectivity between 0 and 1. Let $g_{A,c}(a,b)$ be a weighted connectivity function defined as follows:

$$g_{A,c}(a,b) = \begin{cases} \frac{T_{a,b}}{T^{sum}} & \text{if clients } a \text{ and } b \text{ are connected} \\ & \text{by a set of non-compromised nodes} \\ 0 & \text{otherwise} \end{cases}$$

The weighted expected client connectivity is (same as expected client connectivity from Definition 1 with the exception of using function g instead of f):

$$E\left[\sum_{\{a,b \in M : a<b\}} g_{A,c}(a,b)\right] = \sum_{c \in C}\left(\prod_{e_k \in c} P(e_k) \prod_{e_k \notin c}(1 - P(e_k))\right) * \left(\sum_{\{a,b \in M : a<b\}} g_{A,c}(a,b)\right)$$

The Weighted Diversity Assignment Problem is:

$$\text{argmax}_A \left(E\left[\sum_{\{a,b \in M : a<b\}} g_{A,c}(a,b)\right]\right)$$

We know that W-DAP is also hard to solve given that DAP is a special case of W-DAP, and we proved DAP is hard to solve. This special case occurs when all traffic weights are equal.

Theorem 4: The Weighted Diversity Assignment Problem is NP-Hard with two or more variants.

Proof: In the case where $\forall a, b, c, d \in M, T_{a,b} = T_{c,d}$, solving W-DAP is equivalent to DAP. ■

B. MIP Approach to W-DAP

For the MIP formulation we keep the same formulation of DAP from Section III-B while reformulating the objective and certain constraints to correctly include traffic weights between clients. The following modifications to DAP will allow W-DAP to be solved.

W-DAP objective:

$$\text{maximize}_{s,f} \quad \frac{1}{2 * T^{sum}} * \sum_{c \in C, a \in M, x \in N} \left(\prod_{e_i \in c} P(e_i) \prod_{e_i \notin c}(1 - P(e_i))\right) * f_{c,a,a,x} \tag{19}$$

We maximize the weighted expected client connectivity of the graph, over all compromise events. The first normalizing term changes to $\frac{1}{2 * T^{sum}}$ instead of the number of client pairs. This is just a detail that normalizes the solutions of W-DAP

between 0 and 1, and the 2 in the denominator is to handle a small discrepancy where T^{sum} sums over each client-pair once while here we sum over each client pair twice (both directions). The connectivity values $f_{c,a,a,x}$ will be weighted correctly due to the following changes in constraints.

Weighted client flow constraints (I):

$$\sum_{x \in N} f_{c,a,x,b} \leq T_{a,b},\ c \in C,\ a,b \in M,\ a \neq b \tag{20}$$

This is the main change that ensures flow variables optimize for weights. Clients accept an amount of flow according to the traffic weights. This forces the flow from a client a to a client b summed over all paths to take on a value of $T_{a,b}$ if and only if at least one path from a to b exists given the compromise event c.

Weighted topology constraints:

$$f_{c,a,i,j} \leq \sum_{b \in M, a \neq b} (T_{a,b}) * w_{i,j},\ c \in C,\ a \in M,\ i,j \in (N \cup M) \tag{21}$$

Here, $w_{i,j}$ denotes whether an edge exists (value of 1) or does not exist (value of 0) between node i and j. We must change the previous value of $(|M| - 1)$ with this summation over $T_{a,b}$ values since this is the largest possible amount of flow that may be needed for a given edge.

Weighted variant flow constraints (I):

$$f_{c,a,x,i} \leq \sum_{b \in M, a \neq b} (T_{a,b}) * \min_{e_i \in c}(1 - s_{v_i,x}),\ c \in C,\ a \in M,\ x \in N,\ i \in N \cup M \tag{22}$$

Weighted variant flow constraints (II):

$$f_{c,a,i,x} \leq \sum_{b \in M, a \neq b} (T_{a,b}) * \min_{e_i \in c}(1 - s_{v_i,x}),\ c \in C,\ a \in M,\ i \in N \cup M,\ x \in N \tag{23}$$

These two weighted variant flow constraints also change the maximum amount of flow for an edge.

C. W-DAP on the case study topology

We provide the following example of how W-DAP is useful in a realistic scenario. We consider the case with four variants from Section VI. This scenario is selected as optimal assignments must make interesting choices as the topology cannot connect all clients by each of the four variants independently. We let one client-pair be an important client pair such that any possible connectivity is more important than connecting other client-pairs. To do this for the important client pair a, b we let $T_{a,b} = T_{b,a} = 100000$ (exact value unimportant, just needs to be very large) while all other traffic weights are 1. The optimization will treat this client pair a, b as a primary client pair and ensure assignments ensure highest connectivity for that pair. Then, the other client pairs are assigned to maximize connectivity as long as it does not hinder the connectivity between a and b.

Figure 8 illustrates the optimal assignment to W-DAP in this scenario. The weighted expected client connectivity is approximately 0.99925, and this value is significant since it is the probability that at least any one variant is not compromised since $1 - 0.1 * 0.15 * 0.2 * 0.25 = 0.99925$. It is nearly this value since the dominant client pair is connected as long as any single variant exists as you can see the four node-disjoint paths between these two clients. This solution is more difficult than just finding any set of four node-disjoint paths between these two clients as that is an easily solvable problem. Out of all possible node-disjoint paths between these two clients, this one maximizes the expected client connectivity of the remaining nodes. Thus, the network finds the best scenario for this primary client pair, and it can still aim to satisfy the other clients pairs as well. Due to the constraint of having to connect this primary client pair, the general expected client connectivity does suffer as it is 0.99482 compared to the 0.9975 value that was found in Figure 7(c).

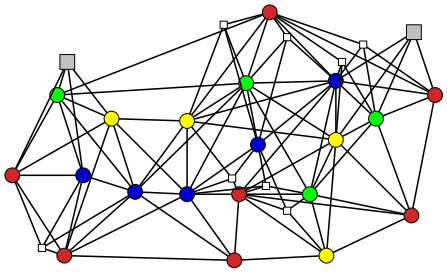


Fig. 8. Assignment with the W-DAP where the traffic weights between two clients (larger squares with gray fill) have a dominant traffic weight (100,000) while all other client pairs have small identical traffic weights (1). The weighted expected client connectivity in this scenario is 0.99925 while the standard expected client connectivity is 0.99482.

VIII. Errors in Compromise Information

Up to now, we have assumed the true assignment compromise values are known and independent with each other. In a realistic scenario, these assignment values could be selected based on expert opinion or extracted from real-world statistics. Both techniques cannot be perfectly accurate. In this section we investigate what occurs when assignment is based on imperfect information.

A. Methodology to investigate erroneous information

We establish certain parameters and values that we use to investigate the effects of errors in information.

We define three scenarios for obtaining an ECC (expected client connectivity) from solving DAP:

- A_ECC_A_INFO is the ECC value based on available information for an assignment solved with the available information. This is the connectivity that a network operator expects when using an assignment based on solving DAP with available information.
- R_ECC_A_INFO is the ECC value based on real information for an assignment solved with available information. This is the realistic connectivity that a network operator will actually achieve when using an assignment based on solving DAP with available information.
- R_ECC_R_INFO is the ECC value based on real information for an assignment solved with real information. This is the connectivity that could have been achieved if the network operator had perfect information.

TABLE III
PROBABILITY DISTRIBUTIONS FOR DIFFERENT α VALUES.

e_3	e_2	e_1	$\alpha = 0$	$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1.0$
			0.612	0.659	0.706	0.800
		1	0.153	0.127	0.102	0.050
	1		0.108	0.094	0.079	0.050
1			0.068	0.051	0.034	0.000
	1	1	0.027	0.020	0.014	0.000
1		1	0.017	0.013	0.009	0.000
1	1		0.012	0.009	0.006	0.000
1	1	1	0.003	0.027	0.052	0.100

We consider two types of discrepancies between available and real information. First, some compromise events have inaccurate values, that is, $P'(e_i) = P(e_i) + \Delta_i$ where P' is the available probability distribution, P is the actual probability distribution, and Δ_i is the error for a particular compromise event. Second, the compromise events are not fully independent, that is, $P'(E) = (1-\alpha) * P(E) + \alpha * D(E)$ where E is a set of compromise events, $D(\cdot)$ is the probability distribution if there is complete dependence among the events, and α is a parameter determining how correlated the variants actually are ($\alpha = 0$ is complete independence while $\alpha = 1$ is the most extreme dependence). We illustrate the independent and full dependence scenarios with Venn diagrams in Figure 9(a). We also show in Table III an example of the probability distribution of $P'(\cdot)$ for differing values of α with three variants $P(e_1) = 0.1, P(e_2) = 0.15, P(e_3) = 0.2$.

With a discrepancy between the available and real information and letting $x =$ A_ECC_A_INFO, $y =$ R_ECC_A_INFO, and $z =$ R_ECC_R_INFO we observe the following two types of errors.

- CONFIDENCE_ERROR $= \frac{|x-y|}{y}$ is the error in how confident a network operator is with the created assignment.
- CONNECTIVITY_ERROR $= \frac{|y-z|}{z}$ is the error in how much worse an assignment based on available information is versus an assignment based on the real information.

B. Error analysis on random topologies

We show the effect of a discrepancy in the compromise probability of a single variant. Then, we show the effect of discrepancy in the assumption of complete independence among variants. We use random topologies with similar settings to the random topologies in Section IV-D. Each topology had 5 clients and 3 variants with compromise probabilities $P(e_1) = 0.1, P(e_2) = 0.15, P(e_3) = 0.2$. In that section we showed results when varying density and number of nodes. For varying density we fixed the number of nodes at 25, and for varying the number of nodes we fixed the density at 6. These values were chosen as these parameters produced interesting topologies, that is, the topologies were connected but not too connected that assignment was trivial. Thus, in this section we fix the number of nodes to 25 and density to 6 for interesting topologies to investigate the effects on assignment when there are discrepancies between available and real information.

Figure 9(b) shows the CONFIDENCE_ERROR and CONNECTIVITY_ERROR when the $P(e_2)$ used for assignment is different from the real $P(e_2)$. We show the errors when the available information has a compromise probability greater than the actual compromise probability ($\Delta_2 < 0$) and less than the actual compromise probability ($\Delta_2 > 0$). We see the greatest errors (for both types) when $P(e_2)$ is believed to be a weaker variant than it truly is, that is, $\Delta_2 < 0$ this is due to the assignment algorithm preferring to select v_3 over v_2 when forced to make a choice between these two. We observe little errors when $\Delta_2 > 0$ which is the case that the available information indicates v_2 is a stronger variant than it actually is. This is due to the random topologies having many client pairs that can be connected by two paths, so it is not so detrimental for the assignment to prefer v_1 over v_2. There is some error which indicates that the preference of v_2 over v_1 is slightly detrimental.

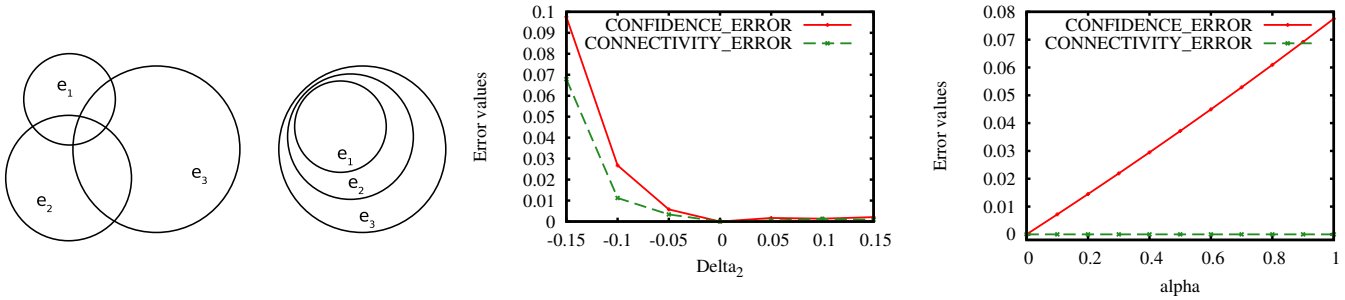
Figure 9(c) shows the errors when the assignment selected is based on the assumption of complete independence. We note in this case that R_ECC_A_INFO = R_ECC_R_INFO, since the assignments are actually the same despite the change in independence information, and thus CONNECTIVITY_ERROR is always equal to zero in this case. However, CONFIDENCE_ERROR is nonzero since any connectivity believed to be found by a network operator is less than the realistic connectivity since dependence among variants is detrimental to diversity. We see that this error increases linearly with α, the parameter controlling dependence.

IX. RELATED WORK

Diversity assignment. The work most similar to ours considers diversity assignment over nodes of a distributed system [18], but the goal of that work is to prevent the spread of malware. In contrast, we assume that if a node of some variant is compromised, then all nodes of that variant are also compromised, as the attacker is not restricted to only using links within the network. To assign diversity to prevent the spread of malware, the computation problem in [18] is different from ours as they intend to minimize the number of links which contain two nodes of the same variant. Thus, their underlying optimization problem for variant assignment is a version of the classic graph coloring algorithm. This problem is NP-Hard, so their work also explores a heuristic solution which can scale to large networks.

Fault-tolerant topology construction. Existing work has introduced the concept of the *fault-diameter* of a graph, which is a metric that bounds the diameter of a graph given that a bounded number of nodes may fail [19], [20], [21], [22]. For a network topology, this means that if the number of failures is bounded, then the maximum number of hops between any two correct nodes will not exceed the fault diameter. This translates to acceptable latency and overhead even in the worst case. Work in this area has considered various ways to create graphs with good fault-diameters, but these methods only consider unweighted graphs where edges are possible between any pair of nodes. In our work, we assume the topology is chosen ahead of time and fixed to ensure good link quality, and we do not need to add edges for our technique.

In wireless contexts, work has studied the allocation of energy among nodes in a wireless adhoc network to ensure high connectivity even when some bounded number of nodes



(a) (b) (c)

Fig. 9. (a) depiction of the difference between independence on the left ($\alpha = 0$) and full dependence on the right ($\alpha = 1$), (b) error values with a discrepancy between real and available information in Δ_2, and (c) error values with a discrepancy between real and available information in α.

fail [23], [24], [25]. The work assumes that node positions are fixed and an amount of energy can be assigned to each node. Higher energy at a node implies a larger transmission range and more possible connections for that node. The optimization problem is to find a power assignment to nodes which minimizes the global power consumption while ensuring connectivity among correct nodes given a bounded number of nodes can fail. This optimization problem is studied in detail, providing a MIP and exploring various approximation techniques.

WSN key distribution. Wireless Sensor Networks (WSNs) consist of resource constrained devices which sense physical phenomena and deliver this information over a wireless newtork to a base station. In this context, PKI and full pairwise key initialization are prohibitive due to the limitations of sensors. Thus, various work proposes special key distributions, where secret information is shared among more than a single pair of nodes [26], [27], [28], [29], [30]. This has similarities to diversity assignment as the physical capture of a single node allows an attacker to utilize the secret information on that node to attack links of other nodes which share similar secret information. Our work does fundamentally differ as we perform diversity assignment with the complete topology information to maximize a resiliency metric while WSN key distribution work focuses on assigning initial secret information to nodes to maximize the potential of many links are secure. With the potential for many secure links, a random wireless topology can be created and have certain resiliency properties.

Path diversity. Other work has studied the possible geographically diverse paths of real-world topologies [31]. The assumptions of this work are that problems on today's Internet are correlated geographically, so having multiple paths which contain nodes that are geographically diverse will result in higher reliability. The main contributions of this work are defining the metric of geographic diversity for a graph and analyzing this value for realistic graphs. No assignment problem exists in this context as diversity is fixed by geographic location.

X. Conclusion

This work illustrates the resiliency benefits gained when shifting from homogeneous networks with potential vulnerabilities shared across all routing nodes to networks that leverage optimally-assigned diversity. We summarize our key findings. First, randomly assigning diversity to a realistic network has surprisingly poor results, which motivated the need to formulate and solve the Diversity Assignment Problem (DAP). Second, we propose an algorithm that solves DAP optimally, and show the results on medium-sized random networks as well as a realistic network. Third, we propose an algorithm that approximates the optimal solution, scaling well to large networks, and show that on random networks, the resulting resiliency is close to that of the optimal solution. Fourth, we show how to optimize for the specific resiliency needs of an application running on the network. We applied this to Paxos and BFT, finding that the probability of making progress can be significantly increased. Lastly, as it is difficult to exactly estimate compromise probabilities we showed how discrepancies between compromise probabilities used for assignment and the real compromise probabilities affect assignment and resilience.

Acknowledgement

This work was supported in part by DARPA grant N660001-1-2-4014. Its contents are solely the responsibility of the authors and do not represent the official view of DARPA or the Department of Defense.

References

[1] M. Garcia, A. Bessani, I. Gashi, N. Neves, and R. Obelheiro, "Os diversity for intrusion tolerance: Myth or reality?" in *Proceedings of DSN*, 2011, pp. 383–394.

[2] B. Cox, D. Evans, A. Filipi, J. Rowanhill, W. Hu, J. Davidson, J. Knight, A. Nguyen-Tuong, and J. Hiser, *N-variant systems: A secretless framework for security through diversity*. Defense Technical Information Center, 2006.

[3] I. Gashi, P. Popov, and L. Strigini, "Fault tolerance via diversity for off-the-shelf products: A study with sql database servers," *Transactions on Dependable and Secure Computing*, vol. 4, no. 4, pp. 280–294, 2007.

[4] Y. Deswarte, K. Kanoun, and J. Laprie, "Diversity against accidental and deliberate faults," in *Proceedings of Computer Security, Dependability and Assurance: From Needs to Solutions*, 1998, pp. 171–181.

[5] "LTN global communications," http://www.ltnglobal.com/, accessed: 5/2/2012.

[6] L. Lamport, "The part-time parliament," *ACM Transactions on Computer Systems*, vol. 16, no. 2, pp. 133–169, 1998.

[7] M. Castro and B. Liskov, "Practical byzantine fault tolerance," in *Proceedings of OSDI*, 1999.

[8] A. Schrijver, *Theory of linear and integer programming*. Wiley, 1998.

[9] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, "Introduction to algorithms third edition." pp. 1161–1161, 2009.

[10] "High-performance software for mathematical programming and optimization," http://www-01.ibm.com/software/integration/optimization/cplex-optimization-studio/, accessed: 5/31/2012.

[11] T. Schouwenaars, B. De Moor, E. Feron, and J. How, "Mixed integer programming for multi-vehicle path planning," in *Proceedings of European Control Conference*, 2001, pp. 2603–2608.

[12] L. Pallottino, E. Feron, and A. Bicchi, "Conflict resolution problems for air traffic management systems solved with mixed integer programming," *Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 3–11, 2002.

[13] G. Huang, B. Baetz, and G. Patry, "Grey integer programming: an application to waste management planning under uncertainty," *European Journal of Operational Research*, vol. 83, no. 3, pp. 594–620, 1995.

[14] "Coin-or," http://www.coin-or.org/.

[15] "Scip," http://scip.zib.de/.

[16] Y. Amir, B. Coan, J. Kirsch, and J. Lane, "Prime: Byzantine replication under attack," *Dependable and Secure Computing*, vol. 8, no. 4, pp. 564–577, 2011.

[17] A. Clement, E. Wong, L. Alvisi, M. Dahlin, and M. Marchetti, "Making byzantine fault tolerant systems tolerate byzantine faults," in *Proceedings of USENIX NSDI*, 2009, pp. 153–168.

[18] A. O'Donnell and H. Sethu, "On achieving software diversity for improved network security using distributed coloring algorithms," in *Proceedings of computer and communications security*, 2004, pp. 121–131.

[19] M. Krishnamoorthy and B. Krishnamurthy, "Fault diameter of interconnection networks," *Computers & Mathematics with Applications*, vol. 13, no. 5, pp. 577–582, 1987.

[20] S. Latifi, "On the fault-diameter of the star graph," *Information Processing Letters*, vol. 46, no. 3, pp. 143–150, 1993.

[21] S. Latifi, "Combinatorial analysis of the fault-diameter of the n-cube," *Transactions on Computers*, vol. 42, no. 1, pp. 27–33, 1993.

[22] K. Day and A. Al-Ayyoub, "Fault diameter of k-ary n-cube networks," *Transactions on Parallel and Distributed Systems*, vol. 8, no. 9, pp. 903–907, 1997.

[23] M. Hajiaghayi, N. Immorlica, and V. Mirrokni, "Power optimization in fault-tolerant topology control algorithms for wireless multi-hop networks," in *Proceedings of MobiCom*, 2003, pp. 300–312.

[24] X. Jia, D. Kim, S. Makki, P. Wan, and C. Yi, "Power assignment for k-connectivity in wireless ad hoc networks," *Journal of Combinatorial Optimization*, vol. 9, no. 2, pp. 213–222, 2005.

[25] D. Panigrahi, P. Duttat, S. Jaiswal, K. Naidu, and R. Rastogi, "Minimum cost topology construction for rural wireless mesh networks," in *Proceedings of INFOCOM*, 2008, pp. 771–779.

[26] S. Çamtepe and B. Yener, "Combinatorial design of key distribution mechanisms for wireless sensor networks," *Transactions on Networking*, pp. 293–308, 2007.

[27] L. Oliveira, H. Wong, M. Bern, R. Dahab, and A. Loureiro, "Secleach-a random key distribution solution for securing clustered sensor networks," in *Network Computing and Applications*, 2006, pp. 145–154.

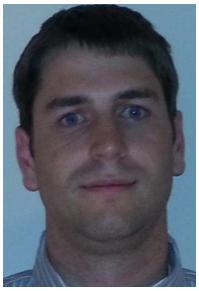
[28] H. Chan, A. Perrig, and D. Song, "Random key predistribution schemes for sensor networks," in *Proceedings of Security and Privacy*, 2003, pp. 197–213.

[29] H. Chan and A. Perrig, "Pike: Peer intermediaries for key establishment in sensor networks," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, 2005, pp. 524–535.

[30] W. Du, J. Deng, Y. Han, P. Varshney, J. Katz, and A. Khalili, "A pairwise key predistribution scheme for wireless sensor networks," *ACM Transactions on Information and System Security*, vol. 8, no. 2, pp. 228–258, 2005.

[31] J. Rohrer, A. Jabbar, and J. Sterbenz, "Path diversification for future internet end-to-end resilience and survivability," *Springer Telecommunication Systems*, 2012.

[32] T. J. Schaefer, "The complexity of satisfiability problems," in *Proceedings of the tenth annual ACM symposium on Theory of computing*, vol. 14, 1978.



Andrew Newell is a PhD candidate in Computer Science at Purdue University. He received his BS in Computer Science and Mathematics at Southern Illinois University at Carbondale in 2008. He is a member of the Dependable and Secure Distributed Systems laboratory. His research interests are in resilient network design, wireless networks, network coding, and machine learning.

Thomas Tantillo is a PhD candidate in Computer Science at the Johns Hopkins University. He received a BS degree in Computer Engineering in 2010 and a MSE degree in Computer Science in 2013 from the Johns Hopkins University. He is a member of the Distributed Systems and Networks laboratory and his research interests include security and intrusion tolerance for networks and distributed systems. He is a member of the IEEE.

Daniel Obenshain is a PhD candidate at the Johns Hopkins University, where he is a Beauchamp Fellow. He received the BS degree from the California Institute of Technology in 2011 and the MSE degree from the Johns Hopkins University in 2013. His research interests include distributed systems and intrusion tolerant systems. He is a member of the ACM and the IEEE.

Cristina Nita-Rotaru is an Associate Professor in the department of Computer Science at Purdue University. She leads the Dependable and Secure Distributed Systems Laboratory. She received BS and MS degrees from Politechnica University of Bucharest, Romania, in 1995 and 1996, and a PhD degree in Computer Science from Johns Hopkins University in 2003. She served on the technical program committee of over 40 conference in networking, distributed systems, and security. She received the NSF CAREER award. Her research interests include security and fault-tolerance for distributed systems and networks. She is a member of the ACM and IEEE Computer Society.

Yair Amir received BS (1985) and MS (1990) degrees from the Technion, Israel Institute of Technology, and a PhD (1995) degree from the Hebrew University of Jerusalem, Israel. He serves as Professor of Computer Science, The Johns Hopkins University since 1995. Prior to his PhD, he gained extensive experience building C3I systems. He is a creator of the Spread and Secure Spread group communication toolkits, the Backhand and Wackamole clustering projects, the Spines overlay network messaging system, and the SMesh wireless mesh network. He has been a member of various program committees including the IEEE International Conference on Distributed Computing Systems, the ACM Conference on Principles of Distributed Computing, and the IEEE/IFIP International Conference on Dependable Systems and Networks. He currently serves as an Associate Editor for the IEEE Transactions on Dependable and Secure Computing. He co-founded Spread Concepts LLC (2000) and LTN Global Communications Inc (2008), and is a member of the ACM and the IEEE Computer Society.

APPENDIX I
PROOF OF THEOREM 1

[1]

Proof: We show that 3-SAT is polynomial-time Turing-reducible to the Diversity Assignment Problem. We will show

that 3-SAT is solvable in polynomial-time if both the DAP is used as a subroutine and the DAP is solvable in polynomial-time.

First we denote the variables for the input boolean expression of the 3-SAT as $\beta_1, \beta_2, ..., \beta_s$. Then, we denote the boolean expression as $(\beta_{\gamma_{1,1}}^{\lambda_{1,1}} + \beta_{\gamma_{1,2}}^{\lambda_{1,2}} + \beta_{\gamma_{1,3}}^{\lambda_{1,3}})(\beta_{\gamma_{2,1}}^{\lambda_{2,1}} + \beta_{\gamma_{2,2}}^{\lambda_{2,2}} + \beta_{\gamma_{2,3}}^{\lambda_{2,3}})...(\beta_{\gamma_{t,1}}^{\lambda_{t,1}} + \beta_{\gamma_{t,2}}^{\lambda_{t,2}} + \beta_{\gamma_{t,3}}^{\lambda_{t,3}})$. For all i and j, $\gamma_{i,j}$ is an index value, so $1 \leq \gamma_{i,j} \leq s$. For all i and j, $\lambda_{i,j} \in \{T, F\}$ where F denotes the complement of the boolean variable while T does not. The 3-SAT problem has s distinct variables and t clauses.

We construct a DAP in the following way to solve 3-SAT. We motivate the intuition for various parts of this construction, but the intuition only becomes correct when all parts are considered together.

Create two variants v_1 and v_2. Create $2s$ nodes denoted by $x_1^T, x_2^T, ..., x_s^T$ and $x_1^F, x_2^F, ...x_s^F$. The problem will be constructed such that for any i, the nodes x_i^T, x_i^F must be assigned variants such that one is v_1 and the other is v_2. With this assignment, x_i^T being assigned v_1 corresponds to β_i being assigned the value true, alternatively x_i^F being assigned v_1 corresponds to β_i is assigned false.

Create $2t$ clients denoted by $a_1, a_2, ..., a_t$ and $b_1, b_2, ..., b_t$. For all i and j add the following two edges $(a_i, x_{\gamma_{i,j}}^{\lambda_{i,j}})$ and $(b_i, x_{\gamma_{i,j}}^{\lambda_{i,j}})$. Each client pair a_i, b_i correspond to a clause in the 3-SAT problem, and the problem will be setup such that at least one of the a_i to b_i paths must have a node assigned with v_1 which will correspond to that node being the one to satisfy this clause.

Create $2{*}s{*}(t{+}1)$ more clients denoted by $a_{1,j}, a_{2,j}, ..., a_{s,j}$ and $b_{1,j}, b_{2,j}, ..., b_{s,j}$ for all j such that $1 \leq j \leq t+1$. For all i and j add the following four edges $(a_{i,j}, x_i^T)$, $(b_{i,j}, x_i^T)$, $(a_{i,j}, x_i^F)$, and $(b_{i,j}, x_i^F)$. Note that for a given i the clients $a_{i,j}$ and $b_{i,j}$ for all j are equivalent in terms of their connections, and each i corresponds to variable in the 3-SAT problem. These clients ensure that every pair of nodes x_i^T, x_i^F is assigned one of each variant which ensures that a boolean variable β_i must be either true or false. The replication of $t{+}1$ client pairs per boolean variable is necessary to ensure that variant choices based on these variables are always more important than those variant choices based on client pairs that correspond to clauses. In terms of the 3-SAT problem, this replication ensures that the assignment of only one value to each boolean variable is never violated even if it helps make many clauses true.

Create nodes denoted by $y_1^T, y_2^T, ..., y_{\left(\binom{|M|}{2} - \frac{|M|}{2}\right)}^T$ and $y_1^F, y_2^F, ..., y_{\left(\binom{|M|}{2} - \frac{|M|}{2}\right)}^F$. These dummy nodes are the most non-trivial part of this construction, but they actually simplify the DAP significantly to ensure variant assignments are meaningful to the 3-SAT problem. All clients have been created in pairs ($\frac{|M|}{2}$ of these pairs), so we want to ensure those pairs meaningful while the other $\left(\binom{|M|}{2} - \frac{|M|}{2}\right)$ client pairs are not interesting. We ensure maximal connectivity between these client pairs that we do not want to be meaningful. To do this we add the following four edges for every client pair a, a' that is not meaningful, (a, y_i^T), (a', y_i^T), (a, y_i^F), and (a', y_i^F) such that a different i is used for each pair a, a'. Thus, trivially, every pair of nodes y_i^T, y_i^F is assigned one of each variant to maximize connectivity.

The last step in the construction is the selection of the compromise event probabilities $P(e_1)$ and $P(e_2)$ along with the minimum expected client connectivity that must be found by the DAP to ensure a 3-SAT solution exists. We consider three types of client pairs which together make up all possible client pairs. First, the $\left(\binom{|M|}{2} - \frac{|M|}{2}\right)$ dummy client pairs should contribute $\frac{\left(\binom{|M|}{2} - \frac{|M|}{2}\right)}{\binom{|M|}{2}} * (1 - P(e_1) * P(e_2))$ to the expected client connectivity value as each pair is trivially connected by both variants. Second, the $s*(t+1)$ client pairs corresponding to 3-SAT boolean variables should contribute $\frac{s*(t+1)}{\binom{|M|}{2}} * (1 - P(e_1) * P(e_2))$ to the expected client connectivity as each pair needs to be connected by both variants. Lastly, the t client pairs corresponding to 3-SAT clauses should contribute $\frac{t}{\binom{|M|}{2}} * (1 - P(e_1))$ to the expected client connectivity as each pair needs to be connected at least by the variant v_1. The assignment of the $P(e_1)$ and $P(e_2)$ values must be such that the expected client connectivity of t client pairs by the variant v_1 is greater than the expected client connectivity of $t-1$ client connected by v_1 and v_2 plus one client pair connected by just v_2. The constraint ensures that a valid to this DAP cannot be $t-1$ clauses that have both true and false variables along with one clause that only has a false variable which is the highest expected value case which is incorrect, so we ensure that the value of this case is always less than the case where all t clauses have just true variables. This constraint is captured by the following inequality $(t-1)(1-P(e_1)P(e_2))+(1-P(e_2)) < t(1-P(e_1))$ which is equivalent to $P(e_1) < \frac{P(e_2)}{(1-t)P(e_2)+t}$. Intuitively, this inequality forces $P(e_1)$ to be sufficiently smaller than $P(e_2)$ to ensure that many connections via nodes with the variant v_2 do not overcome a single connection made by a node with variant v_1. ■

APPENDIX II
PROOF OF THEOREM 3

Proof: We show that a variant of 3-SAT denoted Not-All-Equal 3-SAT [32] is polynomial-time Turing-reducible to CC-DAP. Not-All-Equal 3-SAT has the same setup as 3-SAT except clauses where all variables are true is not allowed; there must be a mixture of true and false variables. We will show that Not-All-Equal 3-SAT is solvable in polynomial-time if both the CC-DAP is used as a subroutine and the CC-DAP is solvable in polynomial-time.

Assume the same network setup as in the proof for NP-Hardness of DAP (Appendix I) which is visualized in Figure 10. This proof differs as we replace the last step of assigning $P(e_1)$ and $P(e_2)$ and use CC-DAP instead of DAP.

In this proof, we can let $P(e_1)$ and $P(e_2)$ take on any value in the range $(0, 1)$ as opposed to requiring certain constraints on these values.

For the CC-DAP algorithm, we aim to maximize the probability of a connected component of $|M|$ clients, i.e., all clients in a connected component.

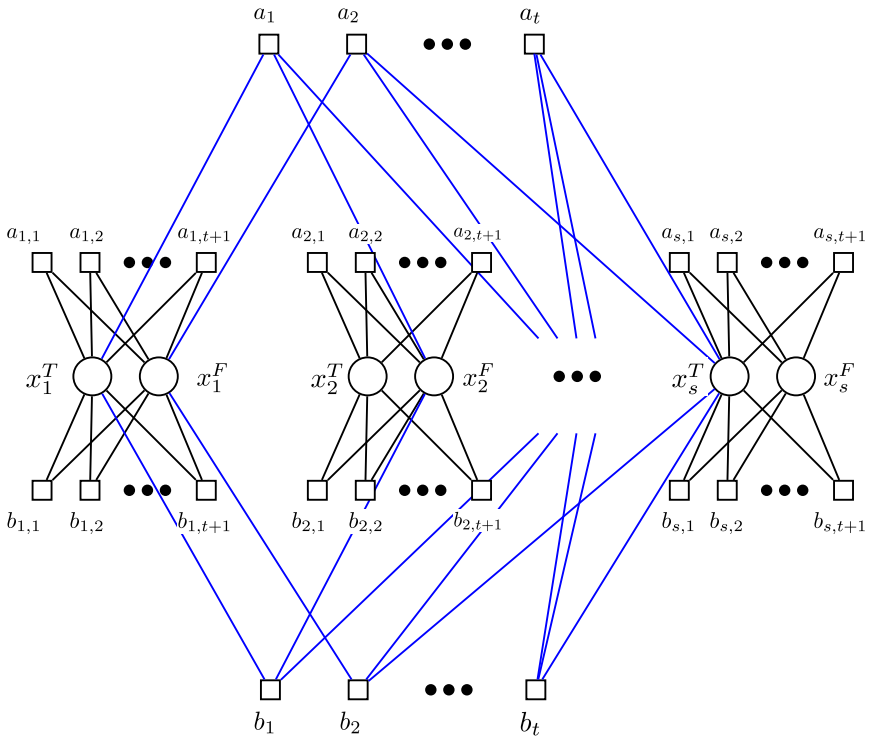


Fig. 10. DAP construction to solve the 3-SAT problem, $(\beta_1 + \bar{\beta_2} + \beta_4)(\bar{\beta_1} + \beta_7 + \beta_s)...(\bar{\beta_8} + \beta_9 + \beta_s)$. Note that the dummy nodes and edges are not included in this diagram.

If and only if CC-DAP finds a probability of $1 - P(e_1) * P(e_2)$ for a connected component of $|M|$ clients, then we have also found a solution to Not-All-Equal 3-SAT due to the following: CC-DAP with a probability of $1 - P(e_1) * P(e_2)$ implies each client pair is connected by both variants v_1 and v_2. The connections between client pairs $a_{i,j}$ and $b_{i,j}$ ensures that $\beta_i \neq \bar{\beta_i}$ for each β_i in Not-All-Equal 3-SAT. The connections between client pairs a_i and b_i ensure that each clause i in the Not-All-Equal 3-SAT problem is connected by at least one true value and at least one false value which is the requirement for Not-All-Equal 3-SAT. Having at least one false value for a clause is a special condition that distinguishes it from standard 3-SAT, and this is the reason we reduce from Not-All-Equal 3-SAT in this proof. ■