CompDrone: Towards Integrated Computational Model and Social Drone Based Wildfire Monitoring

Md Tahmid Rashid, Yang Zhang, Daniel (Yue) Zhang, Dong Wang
Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN, USA
{mrashid, yzhang42, yzhang40, dwang5}@nd.edu

***Abstract*—Forest fires cause irreversible damages worldwide every year. Monitoring wildfire propagation is thus a vital task in mitigating forest fires. While computational model-based wildfire prediction methods provide reasonable accuracy in monitoring wildfire behavior, they are often limited due to the lack of constant availability of real-time meteorological data. In contrast, social-media-driven drone sensing (SDS) is emerging as a new sensing paradigm that detects the early signs of forest fires from online social media feeds and drives the drones for reliable sensing. However, due to the scarcity of social media data in remote regions and limited flight times of drones, SDS solutions often underperform in large-scale forest fires. In this paper, we present CompDrone, a wildfire monitoring framework that exploits the collective strengths of computational wildfire modeling and SDS for reliable wildfire monitoring. Two critical challenges exist to integrate computational modeling and SDS together: i) limited availability of social signals in the regions of a forest fire; and ii) predicting the regions of fire where the drones should be dispatched to. To solve the above challenges, the CompDrone framework leverages techniques from cellular automata, constrained optimization, and game theory. The evaluation results using a real-world wildfire dataset show that CompDrone outperforms the state-of-the-art schemes in effectively predicting wildfire propagation.**

***Index Terms*—social sensing, computational modeling, UAV, wildfire.**

I. Introduction

Wildfires are costly natural phenomena that cause intense ecological and economic impacts globally. An important challenge in wildfire suppression study is the detection of fire propagation, which is necessary to isolate and control wildfire. Computational wildfire models are established techniques that incorporate concrete mathematical theories and meteorological conditions (e.g. wind direction, wind speed, vegetation density) to accurately predict wildfire propagation [1]. However, such models require constant availability of real-time meteorological data for reliable prediction which may not be available in many real-world situations (e.g. inside a forest, it may be impossible to install weather sensors and weather satellites have slow update intervals) [2]. On the other hand, social-media-driven drone sensing (SDS) is emerging as a new integrated sensing paradigm that leverages social media signals to drive drones for reliable sensing [3]. Two recent examples of the SDS paradigm are SocialDrone [3] and SEAD [4]. In contrast to computational models, SDS approaches are more empirical since they rely on real-world observations using physical sensors on drones, and are also more pervasive due to the prevalence of social media (e.g., Twitter, Instagram). However, since SDS approaches utilize drones that are known to have limited flight times, they may be inefficient during large scale wildfires that require continuous scanning of huge areas. Moreover, since SDS approaches primarily rely on social media feeds, they may underperform when the social media data is sparse in remote regions (e.g., inside a forest). In this paper, we explore the opportunity of exploiting the complementary nature of computational modeling and SDS for a reliable wildfire monitoring.

Consider the case of the 2019 Amazon rainforest wildfires that took place in the Amazon region during July 2019. Figure 1(a) shows a geo-tagged tweet about smoke from the fire spreading to the city of Sao Paulo in Brazil, a region that is close to the wildfire (left). The figure also shows the drone footage of the actual wildfire in the state of Mato Grosso (right). Figure 1(b) shows a computational model [5] forecasting the wildfire propagation. Intuitively, if a computational model in the vicinity of the reported event (e.g., “smoke”) on social media could be executed, it could dispatch the drones more effectively to monitor the potential propagation of the fire. While using a computational model that solely relies on a set of initial meteorological data may yield reasonable performance initially, the model may be restricted in consistent prediction due to the lack of real-time data [2]. On the other hand, if we only used an SDS approach, we might not obtain satisfactory performance since the tweets that discuss about the signs of fire may not originate from the actual fire locations and the number of tweets may be insufficient for training a reliable prediction model. Instead, if such tweets could be utilized to trigger an SDS system that works in conjunction with a computational model for deriving probable fire locations, it could have potentially yield more effective wildfire monitoring compared to the standalone approaches. However, a few essential technical challenges exist that need to be addressed before such an integrated wildfire monitoring framework can be established.

The first challenge is the limited availability of social media data in the region of a forest fire. A key limitation of using SDS for wildfire monitoring is data sparsity where social media users contribute a small number of observations [6], [7]. During the onset of a forest fire, people in neighbouring cities

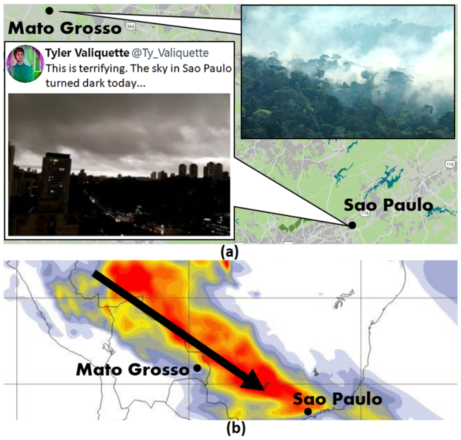

Figure 1. *(a)* A tweet posted during the 2019 Amazon Forest Fire from Sao Paulo (left) and drone footage of the actual fire in Mato Grosso (right). *(b)* A computational model predicting the wildfire propagation.

may report about feeling dry weather or heat, or witnessing smoke. However, while neighbouring urban regions around a wildfire may have Internet connectivity, within the forest itself there may not be adequate cellular coverage or even no Internet connectivity at all. Recent techniques like [8] have attempted to provide internet access in rural areas through a set of balloons, but balloons are difficult to move through a forest. Therefore, it remains an open question on how to develop an efficient mechanism that can leverage the sparse social media data for wildfire monitoring.

The second challenge is identifying the probable regions of fire where the drones should be dispatched to. An SDS-based approach leverages social media signals to dispatch drones to probable locations with fire. However, since drones are expensive resources with limited flight times, during a wildfire encompassing a large forest with several reported locations scattered throughout, it may be incredibly difficult to cover an entire area based just on social media signals. To narrow down scan sites for UAVs, existing autonomous UAV-based systems have: i) integrated satellite-based remote sensing with drones; ii) combined human observations with drones; and iii) fused vehicular network with drones. While these schemes may help in restricting sensing regions for drones in urban areas, they may be limited in the context of a rapidly progressing wildfire, since satellite imagery have slow update intervals (e.g. about once a day), humans have limited mobility, and roadways maybe unavailable. It entails a prediction mechanism that can enhance the scope of sensing for the drones and dispatch them to the right locations.

In this paper, we develop CompDrone, a wildfire monitoring framework that seamlessly integrates computational modelling with SDS-based paradigm to address the above challenges. To address the first challenge, we embed a trigger mechanism that analyzes social media signals in neighbouring regions of a wildfire to drive a data-driven computational wildfire detection model based on cellular automata. The model afterwards fetches meteorological data from an external weather database to predict wildfire propagation. The basic cellular automata model does not consider two important features that affect wildfire propagation: humidity and air quality. We enhance the model by explicitly considering these features for more accurate wildfire prediction. To address the second challenge, the model outputs probable locations of fire that are fed to a game-theoretic task allocation policy for assigning tasks of verifying the predictions to set of drones. This essentially narrows down the sensing scope for the drones as locations with high chances of fire are prioritized. The empirical observations of the drones are then used for augmenting the computational model by updating its key parameters. The CompDrone framework efficiently adapts to the dynamically changing fire conditions while efficiently coordinating the interactions between the computational model and the SDS system. To the best of our knowledge, CompDrone is the first endeavor to combine computational modeling with SDS for integrated wildfire monitoring. The evaluation results with a real-world wildfire dataset show that CompDrone achieves significant performance gains over state-of-the-art wildfire monitoring approaches.

II. Related Work

A. UAV-Based Physical Sensing

Recent advances in physical sensors and drone technology have led to new dimensions in UAV-based physical sensing [9]–[11]. Several studies have proposed UAVs fitted with physical sensors for: i) vehicle identification [9], ii) toxic emission detection [10], iii) disaster damage assessment [12], iv) city-scale video surveillance [11], and v) autonomous forest fire monitoring [13]. However, it may be infeasible to monitor large scale wildfires with a limited number of drones, since the chance of discovering fire is subject to drones patrolling in close vicinity of the fire. In a large forest, launching a limited set of drones that rely on the signals from human operators to be narrowed down to event locations may slow the wildfire monitoring. In contrast, the CompDrone framework melds computational modeling and social media signals to extract probable regions of fire and send UAVs to only those locations, effectively reducing the travel of the UAVs.

B. Social Sensing

Social sensing is rapidly progressing as a pervasive sensing paradigm where humans are used as sensors to obtain information about the physical world [14]–[18]. For example, Boulton *et al.* explored a social sensing-based approach to detect and locate wildfires [19]. Xu *et al.* proposed a framework that carries out semantic and spatial analysis of urban emergency events using social media data [20]. Zhang *et al.* developed a constraint-aware truth discovery model to detect dynamically evolving truth in social sensing [21], [22]. Gu *et al.* presented a predictive heuristic system for augmenting social sensing with data inference [23]. One major limitation of such social sensing schemes is that they solely rely on the noisy social media data that may lead to false alarms in a wildfire and mislead rescue teams to areas with no actual fire [24], [25]. In addition, social signals may be completely unavailable inside a forest with no Internet connectivity where the fire has occurred. As opposed to the above schemes, CompDrone combines SDS-based sensing with computational modelling to address the data reliability and sparsity issues of social sensing.

C. Wildfire Modelling Systems

Computational wildfire modelling systems that incorporate a combination of concrete mathematical theory and physical parameters are being increasingly used in tracking forest fire outbreaks [1], [26], [27]. Filippi *et al.* presented a comprehensive study of several wildfire modelling techniques [1]. A few notable examples of computational wildfire modelling are: i) wavelet-based fire propagation detection [28], ii) cellular automata-based stochastic wildfire modelling [27], and iii) machine learning-based wildfire ignition discovery [26]. A key drawback of existing wildfire modeling techniques is that they are not triggered by real-world observations before the onset of a large wildfire. Only after a fire outbreak when the news of fire is disseminated, wildfire models are run by human operators. Moreover, current wildfire modeling approaches do not incorporate empirical observations from the fire propagating in real-time. In contrast to the above approaches, the CompDrone uses social media signals to initiate the model and utilizes the observations obtained by UAVs to augment the wildfire model for better prediction.

III. Problem Definition

In this section, we formally present the wildfire propagation detection problem for our computational model-driven social drone sensing (CSDS) scenario. We start by defining the key terms and underlying assumptions used in our system. We then introduce the main objective of our problem.

A. Terms and Assumptions

For a CSDS application, we examine a sensing region of interest (ROI) within a given sensing timeline. The sensing timeline is discretized into T periodic intervals, of size Δ each, namely *sensing cycles*. In particular, $t \in [1, T]$ indicates the t^{th} sensing cycle. The sensing ROI is spatially distributed into H separate square *sensing cells*, of length L each, where each cell represents a real-world location. In particular, $SC_{t,h}$ indicates the h^{th} cell at the sensing cycle t.

We also define an essential attribute of the sensing cells, the *cell state*:

DEFINITION 1. Cell State $J_{t,h}$: a variable to indicate the current condition of sensing cell $SC_{t,h}$. A sensing cell is described by one of three possible cell states: I) currently burning; II) not burning; and III) completely burnt.

In our problem, we consider a computational model for determining wildfire propagation, denoted by $\mathcal{Y}$. We use $\widehat{J_{t,h}}$ to indicate the estimated cell state of $SC_{t,h}$ by our CompDrone system. We anticipate that social media users present different *events* on social media which is described as:

DEFINITION 2. Social Media Event $V_{t,p}$: a Boolean variable to denote the existence of noticeable occurrences during the wildfire. In particular, three types of events are considered: i) the visibility of smoke, ii) the weather being dry, and iii) strong wind blowing in a particular direction. In sensing cycle t, the p^{th} event is given by $V_{t,p}$ with a total of N_t events, $\mathcal{V}_t = (V_{t,1}, V_{t,2}, ..., V_{t,N_t})$. We express event $V_{t,p}$ as:

$$V_{t,p} = \begin{cases} 1, & \textit{the event exists in reality} \\ 0, & \textit{the event is reported falsely} \end{cases} \tag{1}$$

Each event is associated with an *event deadline* defined as:

DEFINITION 3. Event Deadline $\delta_{t,p}$: each event is assigned a deadline indicating the level of importance by analyzing spatiotemporal sentiments of the users [29]. For example, an event indicating a city shrouded with smoke may have a tighter deadline than one discussing about dry weather.

The input data from the social media posts (e.g., tweets from Twitter) are denoted by:

DEFINITION 4. Social Signals $\mathcal{S}$: the "signals" contributed by social media users with unknown reliability that represents the early signs of a wildfire.

In our problem, we consider a set of M drones, $\mathcal{D} = (D_1, D_2, ..., D_M)$, each element representing the current location and destination of a drone. We assume that as a drone passes a sensing cell, it collects the status of that cell. We distribute a set of *tasks* to the drones for investigating the state of the cells which is defined as:

DEFINITION 5. Task $W_{t,q}$: a task points to the location of a sensing cell where a drone should be sent out for examining the state of the cell using its onboard sensors. At sensing cycle t, a total of L_t tasks are defined: $\mathcal{W}_t = (W_{t,1}, W_{t,2}, ..., W_{t,L_t})$. In particular, the q^{th} task in $\mathcal{W}_t$ is given by $W_{t,q}$.

Figure 2 shows a scenario for the concepts defined above.

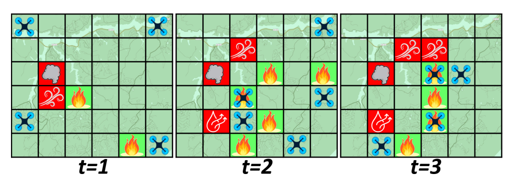

Figure 2. A summary of the sensing cells along sensing cycles. i) The blue drone icons representing the drones; ii) the fire icons mark the sensing cells currently burning (State I); and iii) the red icons indicate the reported events in the social media (e.g. visible smoke, wind conditions, and dryness).

B. Objective

Based on the social signals $\mathcal{S}$, the sensing cells SC_{th}, and the events $\mathcal{V}_t$ with their deadlines $\delta_{t,n}$, the CompDrone framework's objective is to: i) narrow down fire regions from social signals $\mathcal{S}$; ii) identify the cells with probable fire occurrences using the computational model $\mathcal{S}$; and iii) dispatch drones $\mathcal{D}$ to the cells to obtain actual cell states. The eventual goal is to accurately predict the states of the cells using the combination of social signals, drones, and computational modeling to minimize the discrepancy between the estimated cell states and their ground truth. Formally, we solve a constrained optimization problem:

$$\underset{\widehat{J_{t,h}}}{\arg\min} \sum_{n=1}^{M_t} (\mathit{diff}(\widehat{J_{t,h}}, J_{t,h}) | \mathcal{D}, \mathcal{S}, \mathcal{Y}, \mathcal{V}_t, \delta_{t,n}, SC_{t,h}), \quad \forall 1 \le t \le T,\ \forall 1 \le h \le H \tag{2}$$

where *diff* denotes a function that outputs 0 if the values of the two input variables are the same and outputs 1 otherwise.

IV. Solution

In this section, we present the CompDrone framework to solve the wildfire propagation detection problem. The CompDrone is a wildfire monitoring framework incorporating four key modules: i) a Social Signal Distillation (SSD) module that analyzes the unreliable social media data to obtain reliable event reports; ii) a Wildfire Propagation Prediction (WPP) module that uses the event reports from the SSD module as triggers to predict the probable fire regions at each sensing cycle; iii) a Drone Task Assignment (DTA) module that assigns a set of drones to investigate the probable fire regions obtained from the WPP module; and iv) a Parameter Optimization (PO) module that utilizes the measurements obtained by the drones dispatched by the DTA module to augment the fire prediction by the WPP module. An overview of CompDrone is illustrated in Figure 3. We elaborate each module below.

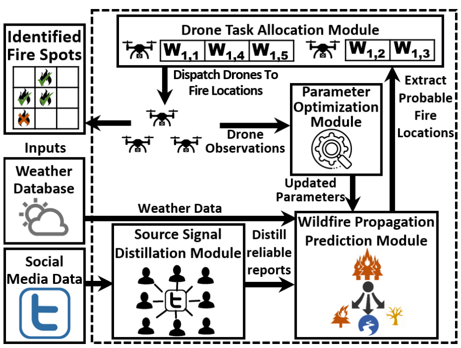

Figure 3. Architecture of the CompDrone Framework

A. Social Signal Distillation (SSD) Module

The SSD module's purpose is to analyze a set of uncertain social media posts with the aim to identify the veracity of reported events during the course of the wildfire [22]. We employ a truth discovery mechanism defined in [3] to obtain reliable signals from noisy social media input data $\mathcal{S}$. Since the social signals are used to determine the early signs of the wildfire, three particular classes of reports are extracted from the social media data: i) abnormal wind conditions, ii) extreme dryness (low humidity), iii) and visible smoke (poor air quality). The output of the truth discovery solution used in our framework is represented by:

DEFINITION 6. Event Correctness $Corr_{t,p}$**:** A measure of the veracity of a reported event $V_{t,p}$ represented by a score in the range (0,1]. Intuitively, a larger value of $Corr_{t,p}$ signifies that event $V_{t,p}$ has a greater chance of truly existing.

B. Wildfire Propagation Prediction (WPP) Module

The WPP module determines the probable fire regions using the reports obtained from the SSD module as triggers. We incorporate an established computational wildfire propagation model that is based on cellular automata [27]. The rationale for our selection is that the model can simulate the fire spread while considering the wind conditions, vegetation conditions and land topography, all of which are highly dynamic during large wildfires. We note that basic cellular automata model in [27] does not consider the weather effects caused by humidity and air quality. We develop our model by explicitly considering both the factors in our solution.

When the event correctness for a set of reported events in a certain region exceeds a predetermined threshold (i.e., $Corr_{t,p} > Corr_{thresh}$), the cellular automata model is initialized. In Section V we discuss how the threshold $Corr_{thresh}$ is determined. Initially, a set of meteorological data (i.e. humidity, air quality, and wind conditions) is fetched from an external database [30] and fed to the cellular automata model. Afterwards, the model uses actual meteorological measurements obtained by a set of drones discussed in the DTA module. The model follows a collection of rules to determine the state of the sensing cells as explained below.

1) State Transition Rules: The rules of evolution for the model relies on the humidity, wind conditions, vegetation, air quality, and slope (i.e. ground elevation). We define $p_{t,h,burn}$ to indicate the probability with which sensing cell SC_{th} burns. $p_{t,h,burn}$ is a function of all the factors that affect fire spread [27]. We further define probability p_{thres} to indicate a threshold probability for a sensing cell to be burning. p_{thres} is determined experimentally based on prior fire conditions. The state of a sensing cell evolves based on the rules expressed by the state diagram in Figure 4.

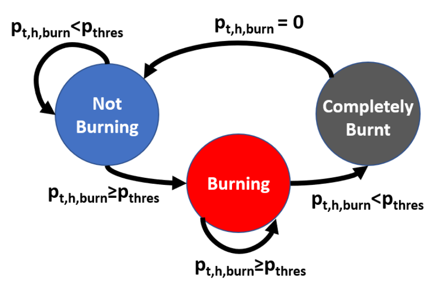

Figure 4. State Transition Rules for Fire Propagation

At the end of a sensing cycle, the rules for transitioning to the next state is decided by observing the present state and the burn probability $p_{t,h,burn}$. If a cell is burning in sensing cycle t and $p_{t,h,burn} \ge p_{thres}$, it will continue to burn in sensing cycle $t+1$. However, if a burning cell has $p_{t,h,burn} < p_{thres}$, it will be completely burned in $t+1$ and $p_{t,h,burn} = 0$ for future sensing cycles. If a cell is not currently burning, but $p_{t,h,burn} \ge p_{thres}$, it will start to burn in $t+1$. Finally, if a cell is not burning and $p_{t,h,burn} < p_{thres}$, it will not burn in $t+1$. Formally, we define $p_{t,h,burn}$ as:

$$p_{t,h,burn} = p_{t,0}(1 + p_{t,h,veg})(1 + p_{t,h,den}) p_{t,h,wea} p_{t,h,slp} \tag{3}$$

where $p_{t,0}$ is a reference probability of any cell in the vicinity of a burning cell to start burning at no wind, minimum dryness, healthy air quality, and flat terrain. $p_{t,h,den}$ indicates the fire spread probability due to vegetation density, $p_{t,h,veg}$ indicates the fire spread probability due to vegetation type, $p_{t,h,wea}$ denotes the fire spread probability due to the weather effects, and $p_{t,h,slp}$ signifies the fire spread probability due to the slope effect. We anticipate that the probability of a cell to burn is reliant on the cumulative factors of all the above factors according to the model. Thus, according to the chain rule of joint distribution, we multiply them together along with the reference probability $p_{t,0}$ to compute the value of $p_{t,h,burn}$. The magnitude of $p_{t,0}$ is obtained experimentally based on null wind conditions, minimum dryness, healthy air quality, and flat terrain. The values of $p_{t,h,den}$ and $p_{t,h,veg}$ are obtained from historical satellite data which is elaborated on the following subsection.

A physical property of fire is that it burns exponentially faster as the wind conditions (i.e. wind direction and speed) are stronger in the direction of the fire spread [31]. Moreover, if the wind consists of air that is of good quality and low humidity, a fire tends to spread faster. Using these relations, we determine the probability of fire spread due to weather effects, $p_{t,h,wea}$ by:

$$p_{t,h,wea} = exp[V_h(K_t^1 p_{t,h,hum} + K_t^2 p_{t,h,air}(cos(\Theta_{t,h,wnd}) - 1))] \quad (4)$$

where $p_{t,h,hum}$ is the relative humidity, $p_{t,h,air}$ is the normalized air quality ($PM_{2.5}$), V_h is the wind speed, and $\Theta_{t,h,wnd}$ is the angle between the wind direction and current fire direction. K_t^1 and K_t^2 are two intensity coefficients that are determined by the PO module. Intuitively, if humidity is low, air quality is high, and the wind blows strongly in the direction of the fire spread, the fire spread probability due to weather effects will increase. The probability of fire spread due to slope, $p_{t,h,slp}$ in Equation 3 is derived using a process described in [27].

We further identify a phenomenon called *spotting* common during wildfires where burning fragments are transported by wind to nonadjacent cells in the vicinity of a burning cell [31]. This causes new centroids of fire (i.e. new cells with state I). We define a function η_t to determine the number of contiguous sensing cells with spotting during certain wind conditions:

$$\eta_t = \begin{cases} 0, & V_h < V_{crit} \text{ and } \\ & \Theta_{t,h,wnd} > \Theta_{crit} \\ V_{crit} ln(V_h cos(\Theta_{t,h,wnd})) + V_{crit}, & \text{otherwise} \end{cases} \quad (5)$$

where V_{crit} is the critical wind velocity in Beaufort scale (i.e. strong breeze) and Θ_{crit} is the critical propagation angle, obtained using a method described in [27]. Intuitively, the above rule determines the number of neighboring cells affected by spotting given the appropriate wind conditions.

2) Determining vegetation type and density: The vegetation type and density of a particular region remains roughly consistent throughout years [32]. We derive the vegetation density $p_{t,h,den}$ of a sensing cell from historical observations of the Sentinel-2 satellite data using Google Earth Engine [33]. We compute the Enhanced Vegetation Index (EVI) which indicates vegetation density. Specifically, we use the atmospherically-corrected surface reluctance from the NIR (Band 8), RED (Band 4), and BLUE (Band 2) bands to derive the vegetation density using a method defined in [32]. To derive the vegetation type, $p_{t,h,veg}$ we spatially cluster the mean annual EVI data of the target sensing region into 4 discrete categories: no vegetation, agricultural areas, shrubs, and dense forests [31]. These categories corresponding to different classes of vegetation with corresponding magnitudes. Figure 5 demonstrates a region of interest with its derived vegetation types.

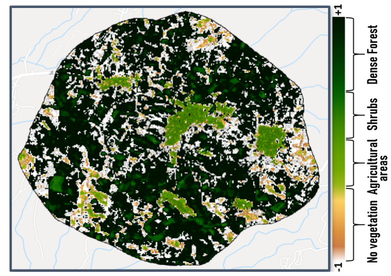

Figure 5. Deriving vegetation type from EVI data

Sensing cells that have high probability of fire (i.e., $p_{t,h,burn} > p_{t,h,thres}$) are investigated using a set of drones as elaborated in the Drone Task Allocation (DTA) module.

C. Drone Task Allocation (DTA) Module

Once the WPP module generates the probability of fire in the sensing cells, the DTA module's purpose is to distribute the tasks of investigating the cells to a set of drones. Due to a limited set of drones, the scheme needs to make a choice consisting of a subset of sensing cells for verification. To accomplish this, we employ a bottom-up game theoretic (BGT) task distribution model as described in [34] to distribute tasks to the drones for exploring the target cells. Furthermore, we use a dynamic routing policy to maneuver the drone to the cells while minimizing the time of travel.

1) Bottom-Up Game-Theoretic (BGT) Task Distribution: A key design philosophy of the BGT task distribution is to model the drones as agents. The agents are allowed to articulate their individual task preference based on the observed payoff of each task. For our model, we determine two key components, the *payoff function* and the *congestion rate*:

DEFINITION 7. Payoff Function $u_{t,h}^m$: a function to denote the individual payoff observed by drone D_m in selecting task $W_{t,q}$. The function sequences the tasks for the drones based several factors: i) uncertainty of the reports computed from the output of the SSD module, $h(Corr_{t,p})$; ii) the distance to the target cell from drone D_m, $\omega_{t,h}^m$ computed using [35]; iii) the event deadline of the cell's event, $\delta_{t,h}$; and iv) the probability of fire in the cell $p_{t,h,burn}$ obtained from the WPP module.

The congestion rate $\gamma_{t,h}$ indicates the level of "crowdedness" of a sensing cell. We compute $\gamma_{t,h}$ using a process described in [34]. Using the above notations, we define the payoff function as:

$$u_{t,h}^m = \frac{\lambda_1 \cdot h(Corr_{t,h}) + \lambda_2 \cdot \omega_{t,n}^m + \lambda_3 \cdot \delta_{t,h} + \lambda_4 \cdot p_{t,h,burn}}{\gamma_{t,h}} \tag{6}$$

where $\lambda_1, \lambda_2, \lambda_3$, and λ_4 are coefficients of the four factors that are obtained using proportional control [4]. Given the above payoff functions, each drone picks a set of tasks that maximizes its own payoff function by reaching a Nash Equilibrium [3], using the *best response dynamics* algorithm discussed in [36]. After the tasks are assigned to all the drones, they are dispatched to investigate the cells where they use their sensors to determine the cell state and the potential to burn (i.e. the vegetation type and density). This information is fed back to the WPP module for parameter optimization.

2) Dynamic Routing Policy: We use a dynamic routing policy to steer the drones to the target cells with predicted fire once all the investigation tasks are distributed to the drones. We reinstate the facts that events have certain deadlines and drones operate on finite energy. Keeping this in consideration, we select a robust route planner algorithm developed using a distributed motion planning formulation [34].

D. Parameter Optimization (PO) Module

The PO module establishes a closed-loop parameter tuning mechanism that leverages the measurements obtained by the drones to enhance the cellular automata model for future sensing cycles. In particular, we incorporate a non-linear optimization for initially obtaining and updating the input parameters of the propagation model (i.e., $p_{t,0}, K_1, K_2$). We encode the parameters in a tuple Q_t and determine their values that minimize the difference between the predicted burn probability of the sensing cells $p_{t,h,burn}$ and the observed burn condition of the sensing cells by the drones given by $\overline{p_{t,h,burn}}$. Formally, the objective is expressed as:

$$\arg\min_{Q_t} \sum_{h=1}^{H} (abs(\overline{p_{t,h,burn}} - p_{t,h,burn})), \quad \forall 1 \le t \le T,\ \forall 1 \le h \le H \tag{7}$$

We use the Downhill Simplex optimization algorithm to update parameters in subsequent sensing cycles. Initially, we set the value of all the parameters to a maximum value of 1. At each sensing cycle, drones are dispatched using the task allocation scheme discussed above. After a series of drone observations $\overline{p_{t,h,burn}}$ are collected for y sensing cycle, the difference between the predicted and observed probabilities, given by $p_{t,dif}$ is computed. y is an adjustable parameter called the update interval which is discussed in Section V. We then update Q_t by applying a non-linear regression on the values of $p_{t,dif}$ up to $p_{t-y,dif}$. At the end of every y sensing cycles, parameters in Q_t are updated.

V. Evaluation

A. Dataset Collection

In order to evaluate the performance of all the schemes, we use a real-world wildfire dataset collected during the 2019 Amazon Rainforest Wildfires that took place in the Amazon region between July and August 2019 [37]. The dataset contains Twitter data feeds collected using the GetOldTweets data collection tool[1] between July 29-August 20, 2019. In addition to that, we also collected hourly air quality $PM_{2.5}$, relative humidity, and wind measurements from monitoring stations in Brazil [30]. The ground truth labels indicating the actual fire propagation are obtained using historic satellite imagery from Planet Labs [38].

B. Experimental Platform and Setup

In our experiment, we conduct a suite of trace-based data-driven emulations to evaluate all the schemes. We conduct the emulations on a reputed drone simulator called ArduPilot SITL that can simulate real-world environments with wildfire occurrences. The experimental setup consists of i) the ArduPilot SITL simulator; ii) a mapping API with Google Maps; and iii) the CompDrone framework. The mapping API connects the ArduPilot SITL simulator with Google Maps and imitates sending drones to real-world coordinates. ArduPilot SITL communicates with the CompDrone framework through UDP requests. Figure 6 illustrates a screenshot of the SITL simulator interfaced with the CompDrone framework.

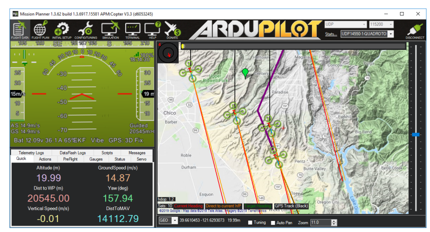

Figure 6. ArduPilot SITL Integrated with CompDrone.

We thereby iterate through the obtained data trace to emulate the wildfire and execute the simulation. We assume that the framework performs data extraction in real-time from the data sources. To obtain the values of the parameters $Corr_{thresh}$, y, λ_1, λ_2, λ_3, and λ_4 in the DTA module, we carry out a learning phase. In particular, we execute a training session in the first $1/3$ of the sensing cycles and set the values of initial parameters which yield the best performance.

C. Baselines Methods

We compare and contrast the following representative baselines in our evaluation which are classified into four categories:

1) Social-Only Systems: In a "social-only" system, the status of wildfire propagation is determined based on the social media posts (i.e., tweets) about fire, wind, smoke, and dryness using social media sensing models. In particular, for our experiments with "social only" baselines, we include two recent social sensing solutions: the Hubs and Authorities (HITS) [39] and Maximum Likelihood Estimation (MLE) [40].

[1]https://pypi.org/project/GetOldTweets3/

2) Drone-Only Systems: In a "drone-only" system, a drone's onboard physical sensors (e.g. thermal scanner and humidity sensor) are used to conduct analysis of a site and identify the states of the cells. We choose two established patrolling strategies used in UAV path planning: i) Random Walk [41] where drones arbitrarily scan a set location and change their courses periodically; and ii) Greedy Heuristics [42] where drones move along designated surveillance routes to detect abnormal occurrences in their paths.

3) Physical Model-Based Systems: Physical model-based systems encompass computational wildfire models that use a combination of mathematical theories and physical parameters to predict the spread of wildfires [1]. We include the cellular automata (CA) model [27] used in our framework as the physical model-based wildfire detection baseline.

4) Social and Drone Combined Systems: In social and drone combined systems, social media signals are used to drive drones to reported locations of the wildfire. For the first baseline, we include a recent social-media-driven drone sensing framework called SEAD [4]. For the second baseline, we create a new scheme by only including the SSD module and the DTA module from the CompDrone framework, namely "SSD+DTA".

D. Experimental Results

1) Detection Effectiveness Across Entire Dataset: For the first set of experiments, we validate all the schemes on the entire dataset and average the results for all sensing cycles. For the drone-based schemes, we set the number of drones to be 50 for the evaluation [2]. The results are presented in Table I. The outcome demonstrates that CompDrone framework significantly outperforms the other baseline approaches in identifying the correct cell status during the wildfire propagation. In terms of the classification accuracy, precision, recall and F1 Score, the performance gains achieved by CompDrone compared to the best-performing baseline, the "SSD+DTA" scheme are 6.3%, 7.1%, 4.8% and 6.7% respectively. We attribute these performance gains primarily to the WPP module which identifies the probable regions of the fire and reduces travels of the drones, helping to cover more cells in a given time period. In addition, the PO module leverages the observations of the drones to enhance the cellular automata model for better fire propagation detection in subsequent sensing cycles.

2) Detection Effectiveness Across Spatial Segmentation: For the second set of experiments, we split the dataset spatially across three different subareas neighbouring the Amazon region to have a closer observation at the detection process, averaging the results for all sensing cycles. The results are presented in Table II. We continue to observe that the CompDrone outperforms the compared baselines. When segmented across subareas, the WPP module in the CompDrone framework is able to execute different and more customized fire propagation models for each area, which leads to better performance in some areas.

[2]The number of drones is selected by considering both the coverage of the sensing area and flight time of the drones.

Table I
OVERALL PERFORMANCE OF WILDFIRE IDENTIFICATION

Category	Algorithm	Accuracy	Precision	Recall	F1-Score
Social Only	HITS	0.320	0.361	0.588	0.444
	MLE	0.514	0.476	0.775	0.590
Drone Only	Greedy Heuristics	0.161	0.191	0.341	0.243
	Rand. Walk	0.121	0.143	0.283	0.188
Social+Drone	SEAD	0.592	0.561	0.802	0.724
	SSD+DTA	0.595	0.564	0.809	0.661
Physical Model	CA	0.583	0.557	0.799	0.654
Our Scheme	**CompDrone**	**0.658**	**0.635**	**0.857**	**0.728**

VI. CONCLUSION

In this paper we present the CompDrone framework, a novel wildfire monitoring framework that integrates computational modeling and social-media-driven drone sensing (SDS). The CompDrone framework mitigates two intrinsic challenges that have not been solved by existing works: i) limited availability of social media data in forest fire regions; and ii) identifying the probable fire locations for dispatching the drones. The evaluation results from a real-world wildfire case study show notable performance gains of CompDrone compared to the baselines in terms of more effective wildfire monitoring.

ACKNOWLEDGMENT

This research is supported in part by the National Science Foundation under Grant No. CNS-1845639, CNS-1831669, Army Research Office under Grant W911NF-17-1-0409. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

REFERENCES

[1] J.-B. Filippi, V. Mallet, and B. Nader, "Representation and evaluation of wildfire propagation simulations," *IJWR*, vol. 23, no. 1, pp. 46–57, 2014.

[2] M. E. Alexander and M. G. Cruz, "Limitations on the accuracy of model predictions of wildland fire behaviour: a state-of-the-knowledge overview," *The forestry chronicle*, vol. 89, no. 3, pp. 372–383, 2013.

[3] M. T. Rashid, D. Zhang, L. Shang, and D. Wang, "An integrated social media and drone sensing system for reliable disaster response," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 2020.

[4] M. T. Rashid, D. Y. Zhang, L. Shang, and D. Wang, "Sead: Towards a social-media-driven energy-aware drone sensing framework," in *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2019, pp. 647–654.

[5] C. L. Reddington, W. T. Morgan, E. Darbyshire, J. Brito, H. Coe, P. Artaxo, C. E. Scott, J. Marsham, and D. V. Spracklen, "Biomass burning aerosol over the amazon: analysis of aircraft, surface and satellite observations using a global aerosol model." *ACP*, 2019.

[6] D. Wang, T. Abdelzaher, and L. Kaplan, *Social sensing: building reliable systems on unreliable data*. Morgan Kaufmann, 2015.

[7] D. Y. Zhang, R. Han, D. Wang, and C. Huang, "On robust truth discovery in sparse social media sensing," in *Big Data (Big Data), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1076–1081.

[8] S. Katikala, "Google project loon," *InSight: Rivier Academic Journal*, vol. 10, no. 2, pp. 1–6, 2014.

Table II
PERFORMANCE OF WILDFIRE IDENTIFICATION ACROSS SUBAREAS

		Sao Paulo				Manaus				Porto Velho			
Category	Algorithm	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Social Only	HITS	0.354	0.416	0.582	0.485	0.235	0.246	0.531	0.337	0.391	0.441	0.685	0.537
	MLE	0.573	0.501	0.799	0.616	0.416	0.428	0.681	0.525	0.536	0.483	0.821	0.608
Drone Only	Greedy Heuristics	0.173	0.189	0.288	0.228	0.118	0.147	0.344	0.206	0.202	0.247	0.411	0.309
	Rand. Walk	0.150	0.167	0.269	0.206	0.090	0.106	0.294	0.156	0.129	0.160	0.297	0.208
Social+Drone	SEAD	0.633	0.632	0.782	0.699	0.521	0.507	0.732	0.599	0.645	0.566	0.924	0.702
	SSD+DTA	0.645	0.647	0.788	0.710	0.537	0.519	0.746	0.612	0.628	0.548	0.925	0.688
Physical Model	CA	0.621	0.625	0.775	0.692	0.509	0.503	0.710	0.589	0.622	0.546	0.916	0.684
Our Scheme	**CompDrone**	**0.689**	**0.684**	**0.820**	**0.746**	**0.615**	**0.617**	**0.820**	**0.704**	**0.667**	**0.601**	**0.927**	**0.729**

[9] A. Gaszczak, T. P. Breckon, and J. Han, "Real-time people and vehicle detection from uav imagery," in *Intelligent Robots and Computer Vision XXVIII: Algorithms and Techniques*, vol. 7878. International Society for Optics and Photonics, 2011, p. 78780B.

[10] K. Daniel, B. Dusza, A. Lewandowski, and C. Wietfeld, "Airshield: A system-of-systems muav remote sensing architecture for disaster response," in *2009 3rd Systems Conference*. IEEE, 2009, pp. 196–200.

[11] A. Trotta, F. D. Andreagiovanni, M. Di Felice, E. Natalizio, and K. R. Chowdhury, "When uavs ride a bus: towards energy-efficient city-scale video surveillance," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 1043–1051.

[12] J. Fernandez Galarreta, N. Kerle, and M. Gerke, "Uav-based urban structural damage assessment using object-based image analysis and semantic reasoning." *Natural Hazards & Earth System Sciences*, vol. 15, no. 6, 2015.

[13] D. W. Casbeer, R. W. Beard, T. W. McLain, S.-M. Li, and R. K. Mehra, "Forest fire monitoring with multiple small uavs," in *Proceedings of the 2005, American Control Conference, 2005.* IEEE, 2005, pp. 3530–3535.

[14] D. Wang, B. K. Szymanski, T. Abdelzaher, H. Ji, and L. Kaplan, "The age of social sensing," *Computer*, vol. 52, no. 1, pp. 36–45, 2019.

[15] D. Zhang, T. Rashid, X. Li, N. Vance, and D. Wang, "Heteroedge: taming the heterogeneity of edge computing system in social sensing," in *Proceedings of the International Conference on Internet of Things Design and Implementation*, 2019, pp. 37–48.

[16] D. Wang, T. Abdelzaher, L. Kaplan, and C. C. Aggarwal, "Recursive fact-finding: A streaming approach to truth estimation in crowdsourcing applications," in *2013 IEEE 33rd International Conference on Distributed Computing Systems*. IEEE, 2013, pp. 530–539.

[17] Y. Zhang, Y. Lu, D. Zhang, L. Shang, and D. Wang, "Risksens: A multi-view learning approach to identifying risky traffic locations in intelligent transportation systems using social and remote sensing," in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 1544–1553.

[18] Y. Zhang, H. Wang, D. Zhang, Y. Lu, and D. Wang, "Riskcast: social sensing based traffic risk forecasting via inductive multi-view learning," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2019, pp. 154–157.

[19] C. A. Boulton, H. Shotton, and H. T. Williams, "Using social media to detect and locate wildfires," in *Tenth International AAAI Conference on Web and Social Media*, 2016.

[20] Z. Xu, H. Zhang, V. Sugumaran, K.-K. R. Choo, L. Mei, and Y. Zhu, "Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, p. 44, 2016.

[21] D. Y. Zhang, D. Wang, and Y. Zhang, "Constraint-aware dynamic truth discovery in big data social media sensing," in *Big Data (Big Data), 2017 IEEE International Conference on*. IEEE, 2017, pp. 57–66.

[22] D. Zhang, D. Wang, N. Vance, Y. Zhang, and S. Mike, "On scalable and robust truth discovery in big data social media sensing applications," *IEEE Transactions on Big Data*, 2018.

[23] S. Gu, C. Pan, H. Liu, S. Li, S. Hu, L. Su, S. Wang, D. Wang, T. Amin, R. Govindan *et al.*, "Data extrapolation in social sensing for disaster response," in *2014 IEEE International Conference on Distributed Computing in Sensor Systems*. IEEE, 2014, pp. 119–126.

[24] D. Wang, L. Kaplan, T. Abdelzaher, and C. C. Aggarwal, "On credibility estimation tradeoffs in assured social sensing," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1026–1037, 2013.

[25] D. Y. Zhang, J. Badilla, Y. Zhang, and D. Wang, "Towards reliable missing truth discovery in online social media sensing applications," in *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 2018, pp. 143–150.

[26] F. X. Catry, F. C. Rego, F. L. Bação, and F. Moreira, "Modeling and mapping wildfire ignition risk in portugal," *International Journal of Wildland Fire*, vol. 18, no. 8, pp. 921–931, 2010.

[27] J. G. Freire and C. C. DaCamara, "Using cellular automata to simulate wildfire propagation and to assist in fire management," *Natural hazards and earth system sciences*, vol. 19, no. 1, pp. 169–179, 2019.

[28] M. A. Finney, *FARSITE, Fire Area Simulator–model development and evaluation*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998, no. 4.

[29] N. Singh, N. Roy, and A. Gangopadhyay, "Analyzing the sentiment of crowd for improving the emergency response services," in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 2018, pp. 1–8.

[30] "Weather api." [Online]. Available: https://openweathermap.org/api

[31] A. Alexandridis, D. Vakalis, C. I. Siettos, and G. V. Bafas, "A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990," *Applied Mathematics and Computation*, vol. 204, no. 1, pp. 191–201, 2008.

[32] O. Dubovyk, T. Landmann, B. F. Erasmus, A. Tewes, and J. Schellberg, "Monitoring vegetation dynamics with medium resolution modis-evi time series at sub-regional scale in southern africa," *IJAEOG*, vol. 38, pp. 175–183, 2015.

[33] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," *Remote sensing of Environment*, vol. 202, pp. 18–27, 2017.

[34] M. T. Rashid, D. Zhang, Z. Liu, H. Lin, and D. Wang, "Collabdrone: A collaborative spatiotemporal-aware drone sensing system driven by social sensing signals," in *28th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2019, pp. 1–9.

[35] M. T. Rashid, M. S. S. Chowdhury, and M. F. Nawal, "Composite indoor localization and positioning system based on wi-fi repeater technology," in *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016, pp. 2130–2133.

[36] M. T. Rashid, D. Y. Zhang, and D. Wang, "Socialcar: A task allocation framework for socialmedia driven vehicular network sensing systems," in *2019 15th International Conference on Mobile Ad-hoc and Sensor Networks (MSN) 2019*. IEEE, 2019.

[37] H. Escobar, "Amazon fires clearly linked to deforestation, scientists say," 2019.

[38] "Daily satellite imagery and insignts." [Online]. Available: https://www.planet.com/

[39] H. Deng, M. R. Lyu, and I. King, "A generalized co-hits algorithm and its application to bipartite graphs," in *Proceedings of the 15th ACM SIGKDD*. ACM, 2009, pp. 239–248.

[40] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proc. ACM/IEEE 11th Int Information Processing in Sensor Networks (IPSN) Conf*, Apr. 2012, pp. 233–244.

[41] D. Orfanus, E. P. de Freitas, and F. Eliassen, "Self-organization as a supporting paradigm for military uav relay networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 804–807, 2016.

[42] S. Waharte and N. Trigoni, "Supporting search and rescue operations with uavs," in *2010 International Conference on Emerging Security Technologies*. IEEE, 2010, pp. 142–147.