cascades in bug bounty programs

T. Maillart*

UC Berkeley, Berkeley, USA

(Dated: April 6, 2016)

I. RESULTS

Scaling bugs as a function of researchers per program:

$$R \sim c^{\beta} \tag{1}$$

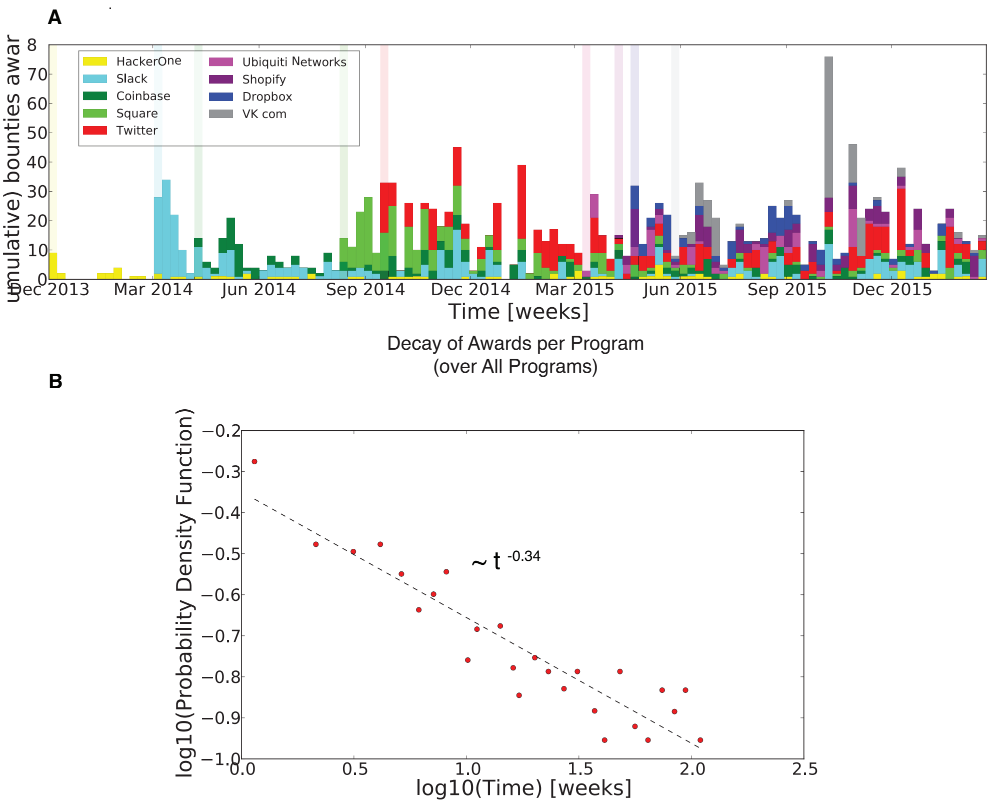

FIG. 1:

*Electronic address: `maillart@berkeley.edu`

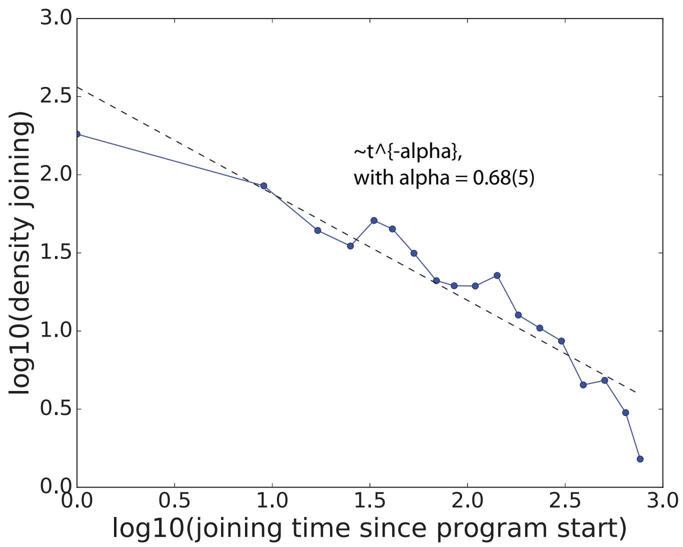

FIG. 2:

with $\beta \approx 1.13$ (see Figure 3).

Let us call $\{R_1, R_2, ..., R_{c-1}, R_c\}$, the total number of bugs found respectively by security researchers $1, 2, ..., c-1, c$. Let us call $R_{\max}(c)$, the largest among the set $\{R_1, R_2, ..., R_{c-1}, R_c\}$. A good estimate of $R_{\max}(c)$ is obtained by the condition that the probability $\int_{R_{\max}(c)}^{+\infty} p(r)dr$ to find a security researcher with a total contribution equal to or larger than $R_{\max}(c)$ times the number c of active developers is equal to 1, i.e., by the definition of $R_{\max}(c)$, there should be typically only one security researcher with such a number of bugs found. This yields

$$R_{\max}(c) \sim c^{1/\mu} \ . \tag{2}$$

An estimate of the typical total number of bugs $R_1 + R_2 + ... + R_c$ identified by the c security researchers can then be obtained as [**? ?**]

$$R_1 + R_2 + ... + R_c \approx c \int_0^{R_{\max}(c)} rp(r)dr \sim c^{1/\mu} \ , \qquad \text{for } \mu < 1 \ . \tag{3}$$

We stress that the scaling $\sim c^{1/\mu}$ only holds for $\mu < 1$ and is replaced by $\sim c$, i.e., linearity, for $\mu > 1$. The upper bound in the integral in (3) reflects that the random variables $\{R_1, R_2, ..., R_{c-1}, R_c\}$ are not larger than $R_{\max}(c)$ by definition of the later. According to equation (3), the typical total production (number of commits) by c developers is proportional to $c^{1/\mu}$, when their contributions are wildly distributed with a power law distribution

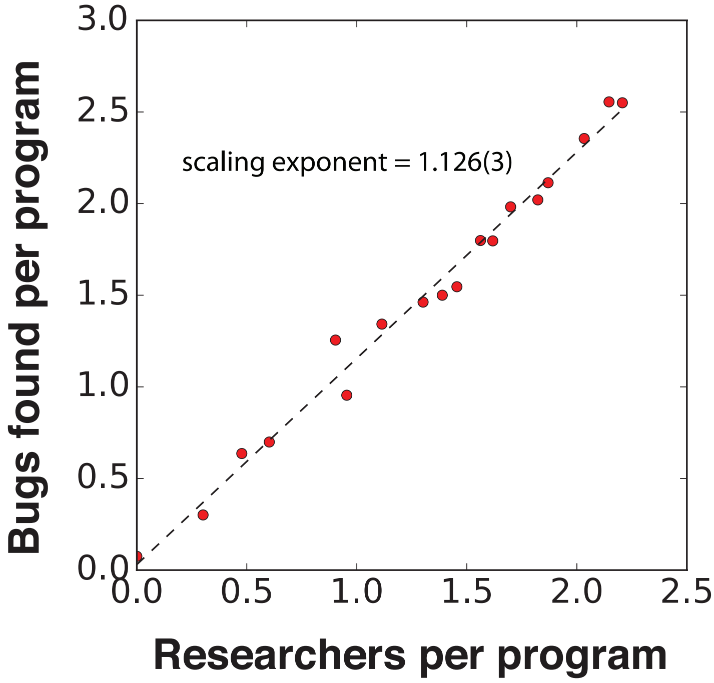

FIG. 3:

with exponent $\mu < 1$. According to this large deviation mechanism, the superlinear exponent β is equal to $1/\mu$.

$$\textbf{prediction of the large deviation mechanism}: \ \beta = 1/\mu \ , \ \text{for} \ \mu < 1 \ . \tag{4}$$

Within this large deviation mechanism, explaining the superlinear productive activity ($\beta > 1$) reduces to explaining the heavy-tailed distribution of commits R per contributor over a large period of time, i.e., amounts to derive the power law distribution (**??**) with $\mu < 1$. For this, the next section proposes a generic model.

$$\textbf{prediction of the large deviation mechanism}: \ \beta = 1/\mu \ , \ \text{for} \ \mu < 1 \ . \tag{5}$$

Distribution of bugs found per programmer per program:

$$P(X > x) = 1/x^{\alpha}, \ with \ \gamma = 1.60(7) \tag{6}$$

Distribution of bugs per program:

$$P_{>}^{tot}(R > r) = 1/x^{\mu}, \ with \ \mu = 0.8(4) \tag{7}$$

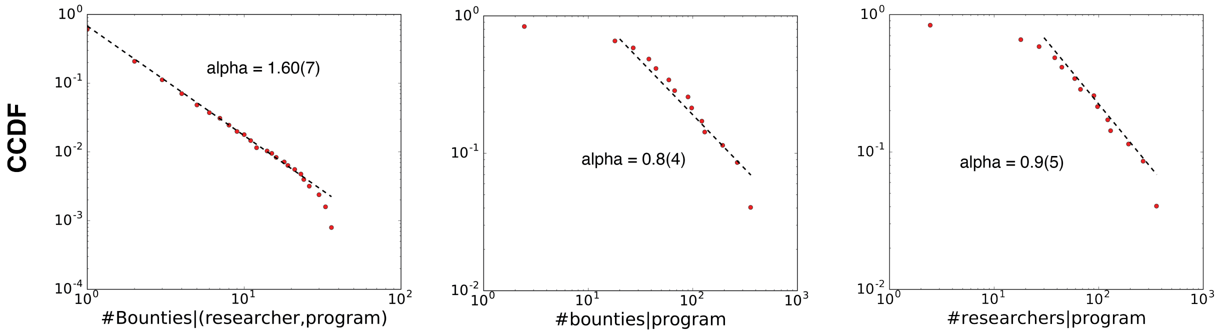

FIG. 4:

Distribution of researchers per program:

$$P(X > x) = 1/x^{\gamma},\ with\ \alpha = 0.9(5) \tag{8}$$

mapping

- superlinear exponent (same)

- bug bounty program $\Leftrightarrow$ OSS contributor