Analyse du système de vote en ligne Neovote

Enka Blanchard, Emmanuel Leblond, Djohar Sidhoum-Rahal, Juliette Walter

Analyse du système de vote en ligne Neovote

Enka Blanchard[1,2] et Emmanuel Leblond [3] et Djohar Sidhoum-Rahal [4] et Juliette Walter [5]

[1] *Laboratoire d'Automatique, de Mécanique et d'Informatique Industrielles et Humaines, UPHF ;* [2] *Centre Internet et Société, CNRS ;* [3] *Scille SAS ;* [4] *Centre de Droit Pénal et de Criminologie, Université Paris Nanterre ;* [5] *Unite Live*

Cet article analyse le système de vote en ligne Neovote, utilisé pour plusieurs scrutins des primaires présidentielles de 2022 (Primaire Populaire, EELV et LR). Nous montrons que les objectifs de transparence, de vérifiabilité et de sécurité exigés par la CNIL et l'ANSSI ne sont pas atteints. Nous montrons l'incohérence du processus de vérification du vote et les vulnérabilités du système qui permettent la publication d'un faux décompte (arrivé en pratique pendant la Primaire Populaire).

Mots-clefs : Cybersécurité, Systèmes de vote, Vote par internet, Étude de cas

1 Introduction

Neovote est l'un des systèmes de vote en ligne les plus utilisés en France, par des institutions publiques comme privées. Indiquant être sélectionné par la CNIL et "homologué" par le Conseil d'État, le Sénat, l'Assemblée nationale, le ministère de l'intérieur et la DGSI, institutions qui ne sont pourtant pas des organismes d'homologation, Neovote ne rend public aucun élément attestant de ces "homologations" [dBGGT22]. Par ailleurs, Neovote a vu ses systèmes de vote remis en cause devant la justice et a été largement critiqué dans les médias, les analyses portant notamment sur la possibilité de s'inscrire plusieurs fois comme électeur.

Neovote n'a pourtant fait l'objet de presque aucune analyse de sécurité de la part de la communauté universitaire (excepté un mémoire de master rendu public alors que nous finissions cet article [dBGGT22]). Notre objectif ici n'est pas de ressasser les éléments déjà critiqués dans la presse mais de faire une analyse indépendante afin de comparer les recommandations juridiques avec la réalité de la transparence et de la vérifiabilité du système mis en oeuvre par Neovote. Nous montrons notamment trois problèmes majeurs :

— ni les propriétés revendiquées par Neovote ni les exigences de la CNIL et de l'ANSSI ne sont atteintes ;
— le système a permis l'affichage temporaire de résultats erronés pendant la Primaire Populaire ;
— le processus de vérification permet à priori ou bien la modification arbitraire de bulletins ou bien la désanonymisation de l'électorat.

Les observations utilisées [†] dans cet article ont toutes été effectuées passivement en documentant le processus de vote sur les ordinateurs de certains co-auteurs inscrits légitimement pour la Primaire Populaire et la primaire EELV, sans chercher à modifier artificiellement les résultats finaux.

2 Recommandations juridiques

La délibération n° 2019-053 du 25 avril 2019 de la CNIL [Com19] est un des documents principaux établissant les normes juridiques du vote en ligne, établissant trois niveaux de risques aux exigences de sécurité croissantes — Neovote annonçant être de niveau 3. Trois exigences concernent notre propos :

— Assurer la transparence de l'urne pour tous les électeurs (objectif n° 2-07).
— Permettre la transparence de l'urne pour tous les électeurs à partir d'outils tiers (objectif n° 3-02).
— S'assurer que le dépouillement de l'urne puisse être vérifié a posteriori (objectif n° 1-11).

†. Un rapport plus détaillé est disponible en pré-version à `http://koliaza.com/neovote-report.pdf`. Nous sommes prêts à partager sur demande la capture vidéo du vote, les fichiers téléchargés (html/js) et les analyses de code.
Nous tenons à remercier Véronique Cortier et Lê Thành Dũng (Tito) Nguyễn pour leurs conseils et leur aide technique.

En suivant le langage universitaire standard (anglophone) des protocoles de vote, on peut assimiler ces exigences à des notions de *vérifiabilité* (comprise ici comme vérifiabilité de bout en bout). Les deux premières correspondent à une combinaison de *cast as intended* et *recorded as cast*, alors qu'on peut assimiler la troisième à *tallied as recorded* [VSD11]. Un détail surprenant est que la transparence du code est un non-dit des systèmes de vote vérifiables. Alors que la transparence (du protocole, du code et des organisateurs) est établie comme nécessaire pour les scrutins en ligne [VSD11] et que la grande majorité des systèmes existant sont open-source, cette propriété n'est pas en elle-même considérée comme strictement nécessaire. Il faut cependant noter que, si le code lui-même n'a pas nécessairement à être ouvert, le protocole devrait l'être. En effet, on peut se demander comment un protocole ayant une boîte noire pourrait être vérifiable (sauf si celle-ci est très localisée ou correspond à un problème à promesse). À défaut d'intégrer l'ouverture du code-source, les recommandations devraient probablement exiger la transparence du protocole (ou des contraintes structurelles sur les entrées/sorties pour toute boîte noire technique).

Le *vade-mecum* de l'Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) répète une partie de ces recommandations dans son Guide de sélection d'algorithmes cryptographiques [Age21]. Se basant sur le principe de Kerckhoffs, il recommande aussi d'éviter de créer des protocoles nouveaux (2.2.6), mais surtout d'éviter de reimplémenter les outils standards (2.2.5) : "C'est pourquoi il est impératif de n'employer que des bibliothèques éprouvées bénéficiant d'un suivi de leur sécurité pour tout appel à des mécanismes cryptographiques."

3 Transparence de Neovote et code extérieur

Le premier point faible que l'on observe dans le système Neovote est que ce dernier n'offre aucune information sur le fonctionnement interne : ni code source, ni documentation ou rapport sur le protocole utilisé, ni les noms des personnes ayant fait le design ou l'implémentation. N'ayant pas accès aux serveurs, les analyses ci-dessous sont donc effectuées en n'utilisant que les informations accessibles côté client.

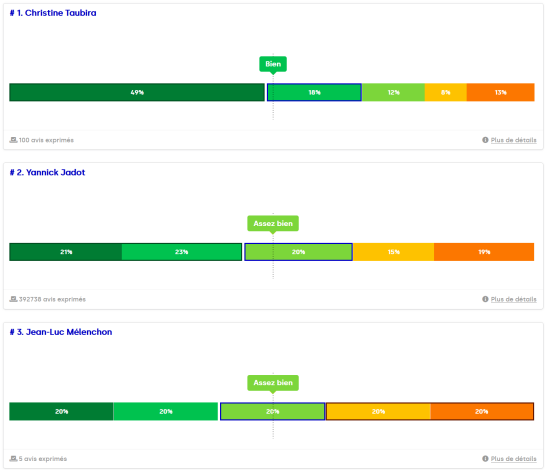
En lien avec ce premier point, trois facteurs compliquent l'analyse et la documentation. Tout d'abord, le code disponible — sous la forme de plusieurs fichiers Javascript appelés dans le html — est partiellement obfusqué, avec les noms de variables et de fonctions subissant apparemment une randomisation à chaque requête. De plus, malgré l'usage de plusieurs OS et navigateurs, nous n'arrivions pas à télécharger directement les fichiers .har (ce qui n'est pas nécessairement lié à une obfuscation volontaire mais limite encore la transparence). Enfin, les sites de Neovote ainsi que celui de la primaire populaire refusaient d'interagir avec la wayback machine (archive.org), limitant la capacité à avoir une copie "neutre" de nos observations.

Un élément mis en avant par Neovote est qu'ils développent tous leurs composants en interne, y compris la pile de sécurité — contrairement aux recommandations de l'ANSSI. Or, nous avons trouvé dans leur code plusieurs exemples de réutilisation du code de la bibliothèque de sécurité asmcrypto.js (disponible sur Github sous licence MIT). Malgré l'obfuscation, on peut par exemple retrouver 32 lignes de code consécutives ayant une structure identique (à remplacement de nom de variables près), correspondant à une fonction utilisée dans le chiffrement AES (AES_Encrypt_process). Le fait qu'ils utilisent une bibliothèque externe pourrait être un point positif, si la celle-ci était appelée lors de la compilation. Nous avons cependant trouvé un autre fragment de code chez Neovote copié depuis une pull-request non intégrée à cette bibliothèque, ce qui indique très probablement une copie manuelle de code, empêchant le suivi et la mise à jour. De plus, cette bibliothèque n'est elle-même plus mise à jour depuis 2018 et a plusieurs protocoles considérés comme obsolescents par l'ANSSI, dont certains se retrouvent dans le code analysé provenant de Neovote. Enfin, cette bibliothèque 1) n'est pas optimisée pour la sécurité mais pour la performance (selon son readme), 2) ne peut pas être considérée comme standard (vu le nombre relativement limité d'utilisateurs/forks et le fait qu'elle ne soit plus maintenue).

4 Vérifiabilité

Le premier problème pour vérifier son vote et l'intégrité du décompte est que Neovote n'offre pas de solution de bout en bout mais communique les résultats aux commanditaires du vote qui se chargent ensuite de sa communication au public. Cela crée un risque d'erreur supplémentaire mais surtout une opportunité pour un adversaire voulant transformer les résultats entre le décompte initial et la publication officielle.

Pour ce qui est de la Primaire Populaire en particulier, la page officielle des résultats le jour du scrutin a temporairement affiché un faux décompte. À cause des pannes fréquentes du serveur de la Primaire Populaire, nous ne savons pas combien de temps exactement cela a duré (au minimum une quinzaine de minutes selon nos observations). Pendant cette durée, les scores des candidats étaient à priori corrects sauf deux exceptions. Premièrement, ceux de Christiane Taubira (dont le prénom était d'ailleurs mal orthographié) pour qui le site n'affichait que 100 suffrages exprimés, suivant les mêmes proportions que ses résultats finaux. Deuxièmement, ceux de Jean-Luc Mélenchon, pour qui le site affichait 5 suffrages (correspondant aux 5 évaluations possibles), comme en atteste la capture d'écran ci-contre, prise peu avant 20h00 le jour du dépouillement sur le site de la Primaire Populaire.



Le deuxième problème est que, selon nos observations et plusieurs observations indépendantes (comme celles de [dBGGT22]), Neovote n'avait pas mis à disposition de procédure de vérification pour le vote de la Primaire Populaire (contrairement aux votes LR et EELV et surtout aux exigences de niveaux 2 et 3 de la CNIL). Des éléments de vérification étaient présents, dont une "preuve de vote" affichée ‡ juste après avoir envoyé son vote. Cependant, aucun mécanisme ne permettait d'utiliser ce reçu : aucune information sur leur site et aucun message envoyé par Neovote ou les organisateurs du scrutin n'indiquait comment utiliser les informations du reçu. Même en se référant au site *www.verifier-mon-vote.fr* utilisé pour d'autres scrutins, aucun des identifiants utilisés lors du scrutin de la Primaire Populaire n'était considéré valide par le serveur. Vu la disparition du système de vérification en ligne — et du code associé [dBGGT22] — le reste de cette section se base sur le code et certaines analyses fournies par un lanceur d'alerte de la primaire EELV (que nous avons pu partiellement authentifier par les nombreux points communs avec le code plus récent).

Un principe fondamental du vote vérifiable est que le reçu ne doit pas permettre de désanonymiser le scrutin. Il doit donc être impossible de prouver comment l'on a voté (sans quoi il y a un risque de coercition et de vente de bulletin). Une parade aux *clash attacks* et *trash attacks* est justement de partager librement les reçus — quoique seulement après le dépouillement [KTV12]. Un point étrange est donc que Neovote prévient ses utilisateurs que le reçu est absolument confidentiel et ne doit pas être partagé.

Le reçu affiché lors du scrutin EELV correspondait à 5 hachés (d'un vote pour un candidat suivi d'une chaîne aléatoire) . L'urne finale (pendant le dépouillement) permet d'associer chaque haché à un vote pour un candidat. Le fait de multiplier les hachés (et potentiellement d'en avoir un pour chaque candidat) permet en apparence de ne pas révéler à quel candidat correspond le reçu. Il y a plusieurs manières de gérer ces hachés côté serveur, mais nos analyses indiquent qu'au moins l'une des conséquences suivantes arrive :

1. les électeurs peuvent indépendamment prouver comment ils ont voté ;
2. le système peut éliminer une grande proportion des votes pour un candidat sans que cela soit visible ;
3. un adversaire ayant accès aux reçus d'un candidat peut en désanonymiser une proportion constante.

L'outil de vérification proposé auparavant par Neovote souffrait d'ailleurs de deux problèmes :

— il est lent sur les grands scrutins et a tendance à renvoyer un échec (à cause du coût de calcul lié au fait que le serveur doive déchiffrer au moins un RSA à 3072 bits par bulletin), facilitant les attaques DDOS.

— étant donné que l'urne n'est pas signée cryptographiquement par Neovote, il est possible de créer une fausse urne donnant un résultat arbitraire en se servant du fichier *extra_hashes.csv*. L'outil de vérification dont nous avons le code n'est pas en mesure de détecter une telle attaque et affiche que l'urne est conforme au reçu.

‡. L'interface n'affichait la preuve qu'une seule fois et son design facilitait le téléchargement par erreur du reçu sans preuve de vote à la place du reçu complet, dont le téléchargement devenait alors impossible.

Enfin, étant donné comment les hachés sont construits lors de la phase de vote, il est facile de créer son propre haché pour le candidat de son choix, puis de le chiffrer avec la même clé que les vrais hachés. Cette "preuve de vote" est indistinguable d'un reçu authentique. N'importe quel électeur peut alors créer un tel reçu, le rendre public et signaler que l'urne a été attaquée. Étant donné que ce signalement n'est pas distinguable d'un signalement authentique, cela remet en question le rôle de ces "preuves de vote".

5 Conclusion

Les éléments développés indiquent donc un nombre important de vulnérabilités et de non-respect des pratiques standards du domaine. Si dans l'absolu de telles insuffisances n'appelleraient qu'une mise au point entre acteurs du domaine, ces problèmes relèvent de l'intérêt général dès lors qu'ils sont le fait du leader du marché qui opère une privatisation de fait de certains mécanismes démocratiques. Face à cette situation, le droit pourrait apparaître tel un régulateur efficace. L'examen des décisions de justice, encore rares, concernant l'activité de Neovote laisse penser que les faiblesses des systèmes de vote électronique ne sont pas encore pleinement saisies par le droit. La Cour administrative d'appel de Marseille a annulé en 2019 un scrutin mis en oeuvre par Neovote en critiquant la procédure de "réassort" qui "n'a pas offert une protection du caractère personnel du vote d'un niveau équivalent à celui des autres modalités de vote" (CAA Marseille, 16 décembre 2019, n°19MA03754). Le reste des faiblesses techniques que nous pointons n'a pas été traité par la cour.

Plus inquiétant, une décision de la Cour de cassation illustre l'insuffisance de la régulation actuelle (Cass. soc., 24 novembre 2021, n°20-17.073). Plusieurs salariés demandaient l'annulation d'un scrutin au motif, notamment, que l'expertise indépendante n'avait pas porté sur le scrutin lui-même ou le code mis en oeuvre, mais sur une version théorique du protocole. La Cour de cassation considère que la seule expertise *in abstracto*, réalisée en amont de la tenue de l'élection, satisfaisait les exigences légales de contrôle. L'expertise réalisée une fois vaut donc pour l'ensemble des scrutins tenus avec ce système et seule une modification substantielle du dit-système commanderait une nouvelle expertise. Or seul un contrôle indépendant à chaque scrutin — par exemple en examinant une signature de code — permet d'examiner si une telle modification a eu lieu (sans pouvoir nécessairement juger de sa substantialité). De plus, certaines attaques ne nécessitent aucun changement de code et certains dysfonctionnements peuvent aussi se produire (comme avec l'affichage du décompte). Enfin, la Cour de cassation ne rejette pas le principe qu'une identification à deux facteurs suffit, même lorsqu'elle est centralisée et permet potentiellement la désanonymisation des électeurs. La dépendance du système au secret d'informations telle que le lieu de naissance, pourtant facilement accessible, ne constitue pas non plus une faille à ses yeux. Ces interprétations sont largement insuffisantes d'un point de vue technique si l'on considère les bonnes pratiques développées par les chercheurs du domaine.

En conclusion, nous n'avons aucune indication que les résultats des scrutins aient été altérés (excepté l'erreur d'affichage). Cependant, les conditions actuelles ne permettent pas de garantir l'intégrité du scrutin et la normalisation de telles pratiques est un risque majeur pour la confiance citoyenne envers les systèmes de vote électronique, qui au surplus pourrait rejaillir sur les outils démocratiques en général. Il apparaît urgent de travailler à une amélioration de la transparence, de la vérifiabilité et du cadre légal de régulation des systèmes de vote électronique.

Références

[Age21] Agence Nationale de la Sécurité des Systèmes d'Information. Guide de sélection d'algorithmes cryptographiques. 2021.

[Com19] Commission Nationale de l'Informatique et des Libertés. Délibération n°2019-053 du 25 avril 2019 portant adoption d'une recommandation relative à la sécurité des systèmes de vote par correspondance électronique, notamment via internet. 2019.

[dBGGT22] Firmin de Barros, Thomas Gergouil, Rémy Grelard, and Samuel Thibault. Analyse de systèmes de vote électronique. Master's thesis, Université de Bordeaux, February 2022.

[KTV12] Ralf Kusters, Tomasz Truderung, and Andreas Vogt. Clash attacks on the verifiability of e-voting systems. In *2012 IEEE Symposium on Security and Privacy*, pages 395–409. IEEE, 2012.

[VSD11] Melanie Volkamer, Oliver Spycher, and Eric Dubuis. Measures to establish trust in internet voting. In *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*, pages 1–10, 2011.